THE ASTROPHYSICAL JOURNAL, 346:566–587, 1989 November 15

A LIMIT ON THE ANISOTROPY OF THE MICROWAVE BACKGROUND RADIATION ON ARC MINUTE SCALES

A. C. S. READHEAD, C. R. LAWRENCE, S. T. MYERS, W. L. W. SARGENT,
H. E. HARDEBECK, AND A. T. MOFFET
Owens Valley Radio Observatory, California Institute of Technology
Received 1988 October 3; accepted 1989 May 15

ABSTRACT

Observations at the Owens Valley Radio Observatory at a frequency of 20 GHz give an upper limit on the anisotropy of the microwave background radiation of $\delta T/T < 1.7 \times 10^{-5}$ (95% confidence) for uncorrelated patches of sky that are uniform on a 2′ scale. This limit is more than a factor of 2 lower than previous limits on comparable angular scales. For the more realistic case of fluctuations with a Gaussian autocorrelation function with coherence angle ϕ_c, the corresponding upper limits are 9.4×10^{-5} ($\phi_c = 12''$), 1.9×10^{-5} ($\phi_c = 2'.6$), and 3.0×10^{-4} ($\phi_c = 25'$). These results place useful constraints on models of galaxy formation based on adiabatic or isocurvature fluctuations in baryonic matter, provided that any reionization of the intergalactic medium occurred at $z < 40$. Adiabatic models are ruled out with greater than 95% confidence, and isocurvature models with $\Omega < 0.8$ are inconsistent with our limits. Theories of galaxy formation that invoke nonbaryonic matter, biased galaxy formation, or a significant fraction of ionized hydrogen at $z > 40$ predict levels of anisotropy a factor of 2 or 3 (and in extreme cases a factor of 10) below the present limit. In the case of nonstandard recombination our limits may provide useful constraints on possible reionization processes. The predictions of most popular contending theories of galaxy formation are within reach of the techniques used in this study.

Subject headings: cosmic background radiation — cosmology — galaxies: formation

I. INTRODUCTION

The origin and nature of the density fluctuations that produced the observed large-scale structure of the universe have been recognized as a fundamental problem since the early days of theoretical cosmology. The pioneering papers by Lemaître (1927) on the "primaeval atom" and of Gamow (1935) on what is now called the "hot big bang" both devoted attention to this problem. The theory of the growth of density fluctuations in an expanding universe was first worked out by Lifshitz (1946). More recently, theoretical activity has been spurred by observations of the isotropy of the microwave background radiation (Partridge 1980*a*, *b*; Uson and Wilkinson 1984*a*, *b*, *c*; Davies *et al.* 1987), the discovery of superclusters and voids (Kirschner *et al.* 1981; Bahcall and Soneira 1982), and indications of large-scale streaming motions (Rubin *et al.* 1976; Aaronson *et al.* 1982, 1986; Bahcall 1987; Dressler *et al.* 1987).

Although several plausible scenarios for galaxy formation have been proposed over the years, ranging from strong perturbation scenarios invoking explosions or shock waves (Doroshkevich, Zel'dovich, and Novikov 1967; Rees 1972; Ostriker and Cowie 1981), to linear, weak perturbation scenarios (Zel'dovich 1967; Harrison 1970; Peebles and Yu 1970), there is no compelling theoretical reason to prefer one over another. As Zel'dovich has pointed out (1972), "No *a priori* preference can be given to small or big perturbation theories—the analysis of observations is the unique approach to the problem."

Observations of the microwave background radiation are important for a number of other reasons, as well:

1. The isotropy of the microwave background radiation, which demonstrates the large-scale homogeneity of the observable universe, is the principle justification for the acceptance of the Robertson-Walker metric and Friedmannian world models.

2. Any structure in the microwave background radiation on a scale larger than $\sim 3°$ must reflect inhomogeneities at decoupling ($z \approx 1500$) and is therefore a probe of the universe at much earlier times than are accessible through direct observations of galaxies and quasars. This may also be true of the structure on smaller scales if there is no early reionization.

3. Fluctuations at decoupling must have derived from physical processes, possibly quantum fluctuations, that occurred in the very early universe. Thus the microwave background radiation is a direct link between the physical processes that gave rise to the fluctuations and the large-scale structure observed today. Indeed, limits on microwave background fluctuations have already proved to be a powerful discriminant between various inflationary scenarios (Hawking 1982; Starobinsky 1982; Guth and Pi 1982; Bardeen, Steinhardt, and Turner 1983).

4. If the intergalactic medium was reionized before $z \approx 40$, the ionized medium will have imposed its own structure on the microwave background radiation on angular scales less than 3°. Background fluctuations on small scales are then a useful fossil record of this important epoch.

Proposed theories of galaxy formation range from those based on adiabatic or entropy fluctuations both with and without nonbaryonic dark matter (Zel'dovich 1967; Peebles and Yu 1970; Wilson and Silk 1981; Vittorio and Silk 1984; Bond and Efstathiou 1987), through those that include early reionization (Hogan 1980; Efstathiou and Bond 1987; Peebles 1987*a*; Vishniac 1987) and biased theories of galaxy formation (Kaiser 1984*a*, *b*; Kaiser 1986), to theories based on explosive galaxy formation (Ikeuchi 1981; Ostriker and Cowie 1981; Ikeuchi, Tomisaka, and Ostriker 1983; Hogan 1984; Vishniac and Ostriker 1985). Cosmic strings have also been invoked as the primary agents of galaxy formation (Zel'dovich 1980; Brandenberger, Albrecht, and Turok 1986; Ostriker,

Thomson, and Witten 1986; Bertschinger and Watts 1988). Expected rms fluctuations in the background temperature range from 3 to 3000 μK. As yet there are no detailed predictions of the angular spectrum of fluctuations produced by strings, although recently some progress has been made on this problem (Ostriker and Thomson 1987; Scherrer 1987; Stebbins 1987; Bouchet, Bennett, and Stebbins 1988).

If intrinsic fluctuations are less than 3 μK, they will be exceedingly difficult to measure in the presence of nonthermal radiation from extragalactic radio sources and the Galaxy (below 30 GHz; Danese, De Zotti, and Mandolesi 1983; Franceschini *et al.* 1988), and thermal radiation from interstellar clouds (above 30 GHz). However, the most popular theories predict rms fluctuations between 15 and 100 μK on angular scales from 1′ to a few degrees, and it is therefore likely that significant progress can be made through observations of the angular spectrum of anisotropies in the microwave background radiation.

In 1982 we installed a sensitive *K*-band maser receiver, built by the Microwave Electronics Group of the Jet Propulsion Laboratory, on the 40 m telescope of the Owens Valley Radio Observatory (OVRO). The broad bandwidth (400 MHz), low noise temperature (26 K), and symmetric feed configuration (separation 7.′15) of this receiver make it suitable for both intrinsic anisotropy observations and for observations of the Sunyaev-Zel'dovich effect in clusters of galaxies. Although atmospheric water vapor fluctuations are a major source of systematic error at centimeter and millimeter wavelengths, the precipitable water above the OVRO is less than 3 mm $\sim$20% of the time from November through March. This cold, arid winter climate makes the OVRO a good site for microwave background observations, and during cold dry periods observations are often limited primarily by the thermal noise of the receiver.

In this paper we report the results of our first five observing epochs. The sensitivity achieved, near 1 part in 10^5, enables us to constrain important cosmological parameters in the context of specific models. We show that a factor of 3 improvement in sensitivity should be possible with this approach in the future, placing almost all theories of galaxy formation within reach.

II. ANTENNA AND RECEIVER

The observations were made on the OVRO 40 m altazimuth telescope. Great care was taken to ensure that the sensitivity of the system was limited only by thermal receiver noise or the atmosphere. This section and the next describe, respectively, how the receiver and the observing strategy were tailored to reduce both systematic and random sources of extraneous noise. Table 1 in § III*a* summarizes the steps that were taken and lists some instrumental problems not mentioned elsewhere that were solved along the way.

Telescope pointing was adjusted at least once every 2 hr by observing a nearby unresolved source. From the size of the required adjustments we estimate the rms pointing error to be 12″. Telescope focus was adjusted automatically as a function of zenith angle, but was checked periodically by the observer.

The maser receiver used in the these observations was based on a design by Moore and Clauss (1979, see also Moore 1980), and has an instantaneous bandwidth of 400 MHz at a center frequency of 20.0 GHz. Figure 1 shows the essential components of the receiver: corrugated scalar feeds (A and B) 3.8 cm apart and symmetric about the telescope axis, which look at areas 7.′15 apart on the sky (1 and 2); a Dicke switch at 10.2 K operating at 10 Hz; the maser itself, with gain G and equivalent noise temperature $T_{\rm maser}$; a second 10 Hz switch, synchronous with the first; two amplifiers with adjustable gains $g_{\rm A}$ and $g_{\rm B}$; and a square law detector that is sampled and digitized at 2 Hz. Dicke-switch transients, which persist for up to 4 ms, are blanked out.

The beam pattern was measured by scanning across 3C 84 in zenith angle and azimuth, over a range of zenith angles from 14° to 69°. The scans were divided into several groups according to zenith angle, and a beam map was produced for each group. While these maps were not identical, the differences were of the same order of magnitude as the level of systematic error in the measurements, judged by the strength of features that were clearly artifacts of the scanning procedure. Accordingly, we added all scans together to achieve the best possible signal-to-noise ratio. The resulting average beam, which should accurately represent the true beam at the pole for all but the lowest contour levels, is shown in Figure 2. As we show in § VIII*c*, the results of this work are quite insensitive to the detailed beam shape at low levels.

A two-dimensional Gaussian fit to the average of the two beams has FWHM of 108″ $\pm$ 1″. The symmetry of the beams ensures cancellation of many possible sources of systematic error such as ground spillover or solar radiation in distant sidelobes. The main beam solid angle measured from this map is $3.2 \pm 0.1 \times 10^{-7}$ sr. The beam and aperture efficiencies, determined from observations of six nonvarying, unresolved objects and the planet Mars, are 0.47 $\pm$ 0.02 and 0.27 $\pm$ 0.01, respectively.

An important feature in the calibration of the observations is that the same noise diode was used in both the determination

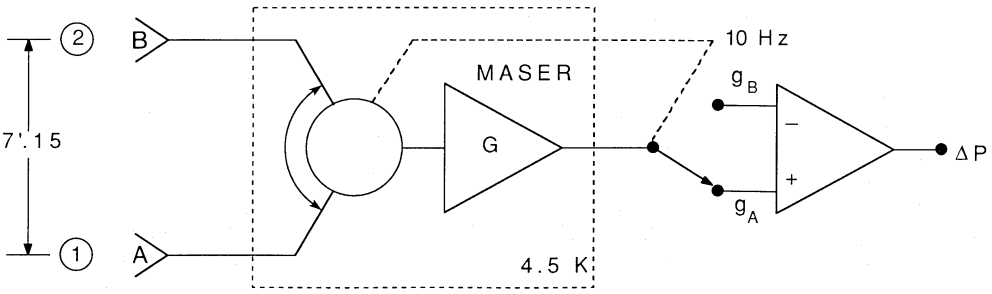

FIG. 1.—Block diagram of the OVRO *K*-band maser receiver. Two feeds, A and B, symmetrical about the radio axis, look at sky areas 1 and 2 separated by 7.′15. The input to the maser (with gain G) is switched between the two feeds at 10 Hz by a cooled switch. The maser output is switched synchronously to amplifiers with adjustable gains $g_{\rm A}$ and $g_{\rm B}$. The output of the radiometer is the power difference between the two sides, and is sampled every 0.5 s. A noise tube is used to calibrate the measurements.

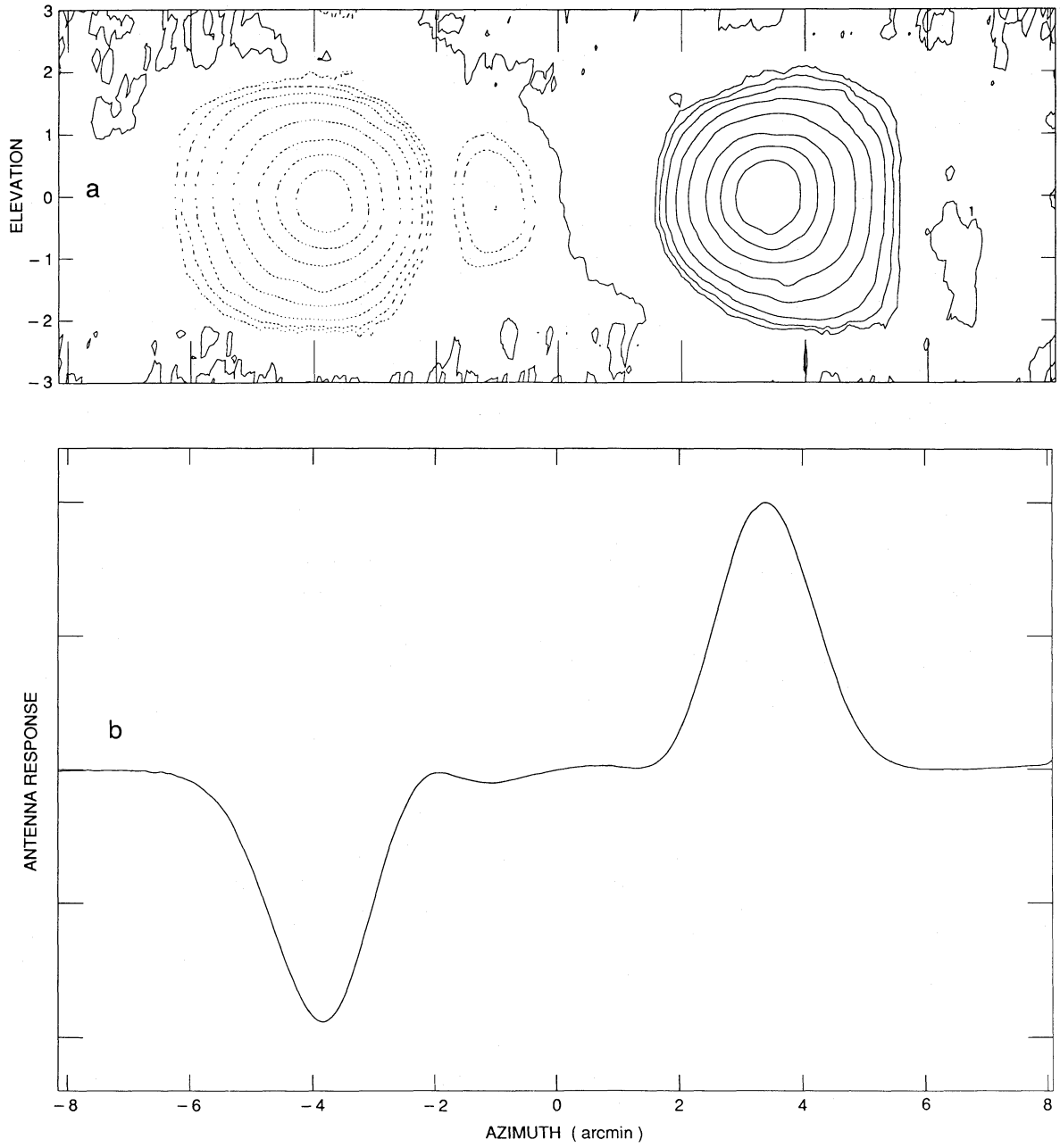

FIG. 2.—(*a*) Contour map of the antenna beam pattern, made from observations of 3C 84 at a zenith angle near 50°. Contour levels are 0, ±2, ±3, ±5, ±10, ±20, ±40, ±60, and ±80% of peak. (*b*) Cross section through the beam map. A two-dimensional Gaussian fit to the average of the two beams has FWHM of 108″ + 1″.

of the beam efficiency and in the calibration of the temperature scale. Any error in the estimated temperature of the noise diode cancels out when the measured temperature differences are divided by the beam efficiency to yield a sky temperature. Thus the only systematic errors in the calibration of our temperature scale arise from the measurement of the beam solid angle and from any variations in the noise diode which last for a significant fraction of the observing epoch.

The total noise temperature of the system at the zenith is typically 40–45 K in good weather, of which 26 K is from the receiver itself, 2.8 K from the microwave background, 4–11 K from the atmosphere, and the rest (presumably) from ground pickup. This gives a thermal noise limit on the sky, corrected for beam efficiency, of 9 $mKs^{1/2}$ for a bandwidth of 400 MHz. The above temperatures were determined from measurements with ambient and 77 K absorbers filling the beams. The noise source used to calibrate the observations was changed several times during the course of the observations reported in this paper. Its temperature was determined initially and then checked regularly either by observations of unresolved sources

(using the flux density scale of Baars *et al.* 1977) or as part of the absolute calibration just described.

Unwanted emission from the ground and the atmosphere, with equivalent noise temperatures T_{gnd} and T_{atm}, enters the feeds along with the interesting radiation T_{sky}. The feeds, wavelengths, and Dicke switch are imperfect, and have losses l_A and l_B in the two arms distributed over components at physical temperatures Θ_A and Θ_B ranging from 4 K to 300 K. The excess noise contributed by these losses can be found by integrating along the two arms: $\int \Theta_A\, dl_A$ and $\int \Theta_B\, dl_B$. The radiometer measures the power difference

$$\begin{aligned}\Delta P = {} & G[g_A(1-l_A)T_{sky1} - g_B(1-l_B)T_{sky2}] \\ & + G[g_A(1-l_A)(T_{atm1}+T_{gndA}) - g_B(1-l_B)(T_{atm2}+T_{gndB})] \\ & + G\left[g_A\int \Theta_A\, dl_A - g_B\int \Theta_B\, dl_B + g_A(1-l_A)T_{inject}\right] \\ & + G[(g_A-g_B)T_{maser}]\,, \end{aligned} \tag{1}$$

where G is the maser gain, T_{maser} is the equivalent noise temperature of the maser, and T_{inject} will be described shortly. Only the first term is astronomically interesting. In an ideal receiver, all other noise sources would be balanced in the two channels, and would cancel out one for one. Specifically, we would like to have (1) $T_{atm1} = T_{atm2}$; (2) $T_{gndA} = T_{gndB}$; (3) $\int \Theta_A\, dl_A = \int \Theta_B\, dl_B$; (4) $l_A = l_B$; and (5) $g_A = g_B$. We will discuss these in turn.

The instantaneous structure of the atmosphere is complicated, and for short integration times the atmospheric terms in equation (1) dominate. However, averaged over a sufficiently long time, T_{atm} will be function of zenith angle only. Accordingly, for observations at constant zenith angle, $\langle T_{atm1}\rangle \approx \langle T_{atm2}\rangle$, where the brackets denote time averages. (The residual effects of atmospheric fluctuations on the data will be discussed in § V.)

Ground pickup and atmospheric effects vary much less with azimuth than with zenith angle, and variations in T_{atm} and T_{gnd} are minimized by moving the telescope as little as possible, particularly in zenith angle. The only part of the sky that can be observed for a long time at almost constant zenith angle is the area near the celestial pole. With an altazimuth mounting, the pole can be observed continuously with only small motions in azimuth as well. Thus by observing near the pole we come as close as possible to satisfying conditions (1) and (2) above.

The construction of the receiver makes significant differences between Θ_A and Θ_B unlikely. Unfortunately, we have no control over the distribution of losses along the two arms, so $\Theta_A = \Theta_B$ does not imply (3) above. In fact, the power measured in the two feeds differs by $\sim$4 K, independent of the absolute power level (e.g., changes in sky temperature). Gain differences in the two horns, for example, would not produce such an effect. This is direct evidence that (3) above does not hold. Therefore, no settings of g_A and g_B can make the last three terms in equation (1) vanish for all values of T_{sky}, T_{atm}, T_{gnd}, Θ, and T_{maser}.

Although we cannot adjust the losses in the two arms, we can inject excess noise ($\sim$4 K) in one arm to balance the *noise* contributions of the losses. In practice, we set $g_A = g_B$, and then adjust T_{inject} for $\Delta P = 0$. The uncompensated difference in *attenuation* due to $l_A \neq l_B$ results in incomplete cancellation of atmospheric and ground noise, but the third and fourth terms in equation (1) are eliminated.

We have found that the value of T_{inject} that gives $\Delta P = 0$ varies diurnally by 0.1–0.4 K. For observations taken as described in § III, this may in the worst case introduce systematic errors of $\sim 10\ \mu$K. While it is clear that in the long run this source of noise must be eliminated, it is too small to affect the observations reported here.

Observing near the NCP, setting the gains equal, and balancing the noise in the premaser components considerably reduce the extent to which instrumental or environmental fluctuations mimic actual temperature differences on the sky. In addition, the 10 Hz switch is faster than most atmospheric fluctuations. Nevertheless, it is not hard to imagine ways in which ΔP could be shifted systematically away from zero for long periods. This problem and its solution are discussed in § III*a*.

III. OBSERVATIONS

a) Strategy

The goal of microwave background isotropy measurements is to detect temperature differences several million times smaller than the equivalent noise temperatures of the equipment used. Even in ideal (thermal noise–limited) conditions, a system with a bandwidth of 400 MHz would require days of integration to reach this level. It is conceivable that for short periods the last three terms of equation (1) could be reduced to the required level using the techniques of § II, but it is inconceivable that this could be done for periods as long as a few days.

The standard solution to this problem is to observe the field being measured for an equal time through both feeds, as shown in Figure 3. We measure

$$\begin{aligned}\mathscr{P} = {} & \frac{1}{\tau}\int_0^{\tau} \Delta P\, dt - \frac{1}{\tau}\int_{\tau+s_1}^{2\tau+s_1} \Delta P'\, dt \\ & - \frac{1}{\tau}\int_{2\tau+s_1}^{3\tau+s_1} \Delta P'\, dt + \frac{1}{\tau}\int_{3\tau+s_1+s_2}^{4\tau+s_1+s_2} \Delta P\, dt\,. \end{aligned} \tag{2}$$

With this switching scheme, all instrumental terms that are constant or vary linearly with time, and all "sky" terms that are constant or vary linearly with position, cancel out if the telescope move times are equal ($s_1 = s_2$).

Instrumental terms quadratic in time survive, but are minimized by choosing τ small; however, since move times s_1 and s_2 are fixed, observing efficiency favors large τ. Most data reported in this paper were taken with $\tau = 20$ s, which provides over 70% observing efficiency and no noticeable increase in noise over $\tau = 10$ s. The telescope control program does not guarantee that $s_1 = s_2$, but move times measured under operating conditions were equal within measurement errors of $\sim$5%.

The combination of Dicke switching between the feeds and antenna switching between sky areas is often called "double switching," and has been used successfully for many observations (e.g., Lake and Partridge 1980; Birkinshaw, Gull, and Hardebeck 1984; Uson and Wilkinson 1984*a*, *b*). Except for atmospheric fluctuations, the largest nonlinear effects that we have thought of would cause errors less than 10^{-5} K even if they persisted for hours.

The temporal and spatial behavior of the atmosphere does have higher order terms that are not removed by double switching. As will be seen in § VI, in the best weather the atmosphere introduces no systematic error, but the total noise

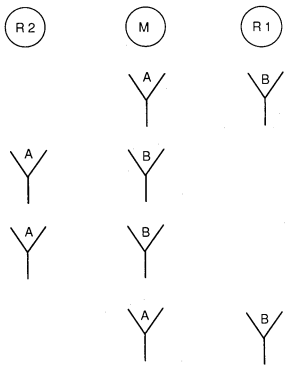

FIG. 3.—Observing cycle. M, R1, and R2 are areas on the sky at the same zenith angle, separated by 7.15. We integrate for time τ, with feeds A and B pointed at M and R1, respectively, then move the antenna in azimuth only (move time = s_1) so that A and B look at R2 and M, respectively. After integration time 2τ, the antenna moves back to the original position (move time = s_2) for rime τ. The cycle then repeats. For each cycle, we measure (see eq. [1])

$$\frac{1}{\tau}\int_0^{\tau}\Delta P\,dt-\frac{1}{\tau}\int_{\tau+s_1}^{2\tau+s_1}\Delta P'\,dt-\frac{1}{\tau}\int_{2\tau+s_1}^{3\tau+s_1}\Delta P'\,dt+\frac{1}{\tau}\int_{3\tau+s_1+s_2}^{4\tau+s_1+s_2}\Delta P\,dt\,,$$

where primed measurements are made with the feeds looking at R2 and M. For one cycle, the temperature difference measured is $\Delta T_i = T_M - \frac{1}{2}(T_{R1} + T_{R2})$.

increases to 30% above the thermal limit. Nevertheless, $\langle T_{atm}\rangle$ is a nonlinear function of zenith angle, and the steady change of zenith angle required to track a source away from the meridian could introduce a sizable systematic error. To minimize this effect, observations must be made not only near the pole, but also at upper or lower culmination. Table 1 summarizes the observing strategy and instrumental adjustments required to control systematic errors.

b) *Observed Fields and the Effective Beam Pattern*

For a fixed total integration time the optimum number of fields to observe for maximum sensitivity depends on the spectrum of microwave background fluctuations. We show in the Appendix that for Gaussian fluctuations, and including the likely effects of systematic errors in the observations, 10–15 fields is best. We chose 12 fields at $\alpha = 1^h, 3^h, \ldots, 23^h$, $\delta = 89°$ (epoch 1985.0), hereafter referred to as NCP 1 through NCP 23. As a check of systematic errors, we initially observed four fields at both upper and lower culmination, and four fields at upper culmination only. Each field was observed for 2 hr centered on transit. The 24 hr sequence consisted of fields NCP 1–NCP 15 at upper culmination, then NCP 5–NCP 11 at lower culmination. Only results for these eight fields are given in this paper, since the number of observations on the remaining four fields is small.

Reference fields R1 and R2 are always at the same zenith angle as M. When a field M is tracked, the sky positions of R1 and R2 change continuously. After 2 hr, the reference fields have become 30° arcs, as shown in Figure 4. These reference arcs are labeled "A1" and "A2" to distinguish them from the simple reference fields R1 and R2 for short observations. Table 2 gives the coordinates of all fields.

The observing cycle shown in Figure 3 usually consists of four 20 s integrations (some early data used 40 s integrations). Since the radiometer sampling time is 0.5 s, each cycle consists of 160 samples. The mean and standard deviation of the samples are recorded. Measurements of the calibrated noise tube every 10 cycles, along with the known beam efficiencies, provide the conversion from power (eq. [1]) to temperature on the sky. Following this conversion, we have for the ith cycle a temperature difference $\Delta T_i \pm \sigma_i$, where

$$\Delta T_i = T_M - \tfrac{1}{2}(T_{R1} + T_{R2})\,. \tag{3}$$

In the analysis to follow error estimates are based on the scatter in the measurements. This will be discussed fully in § IV.

c) *Dates of Observations*

We have undertaken five major series of observations, as shown in Table 3.

TABLE 1

ELIMINATION OF SYSTEMATIC ERRORS

Noise Source	Size	Strategy
Atmosphere: different T_{atm} in two beams	0.1–5 K	Minimize length of track by observing near NCP at upper and lower culmination
		Double switch
		Discard data for which there were clouds within 2 hr
		Discard data for which $\langle T_{atm}\rangle > 11$ K air mass^{-1}, and $(T_{atm})_{rms} > 0.4$ K air mass^{-1}
		Discard data for which $\Delta T_i > 2 \times$ thermal limit
Differential ground spillover	~0.1 K	Minimize length of track by observing near NCP at upper and lower culmination
		Double switch
Receiver:		
Changes in maser gain and noise	0.1–1 K	Dicke switch at 10 Hz
		Set $g_A = g_B$ to better than 0.05%
		Inject noise to balance two channels
Fluctuations in JT return pressure	~0.5 mK	Install large ballast tank in JT return line
Cycling of compressor fans	~20 mK	Stop cycling
Interference (RF pickup on IF cables between telescope and control building)	~0.2 mK	Convert to digital output at telescope
Other	< 50 μK	

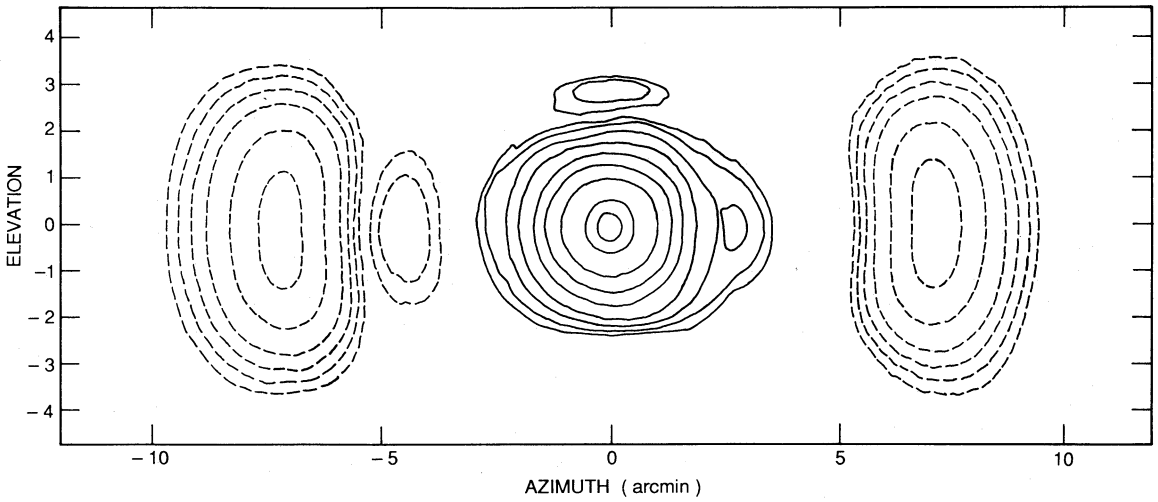

FIG. 4.—Effective beam pattern for 2 hr observations. Contour levels are ±0.5, ±1, ±2, ±4, ±8, ±20, 40, 60, and 80% of peak.

IV. DATA EDITING AND REDUCTION

Bad weather and certain kinds of equipment malfunction produce corrupted data for which there is no statistical remedy. Other instrumental problems produce occasional glitches that can be handled by automatic procedures. In this section we describe how bad data of both types are removed, and how good data are combined.

In general, neither major equipment failures nor bad weather is hard to recognize, but the atmospheric requirements for microwave background anisotropy work at 20 GHz are unusually stringent. It is often not possible to tell by looking at the sky whether conditions are good or bad. Before we turn to the criteria used to reject data, we describe an instrument available for some of the observations that provided an objective measure of atmospheric conditions.

TABLE 2

COORDINATES OF FIELD CENTERS[a]

FIELD	EPOCH 1985.0 R.A.	EPOCH 1985.0 Decl.	EPOCH 1950.0 R.A.	EPOCH 1950.0 Decl.
NCP 1_{A1}	$00^h32^m49^s0$	88°59′34″	$00^h25^m51^s6$	88°47′58″
NCP 1	01 00 00.0	89 00 00	00 48 37.8	88 48 38
NCP 1_{A2}	01 27 11.0	88 59 34	01 11 35.6	88 48 34
NCP 3_{A1}	02 32 49.0	88 59 34	02 07 26.2	88 49 59
NCP 3	03 00 00.0	89 00 00	02 30 42.7	88 51 12
NCP 3_{A2}	03 27 11.0	88 59 34	02 54 34.6	88 51 41
NCP 5_{A1}	04 32 49.0	88 59 34	03 53 03.5	88 54 17
NCP 5	05 00 00.0	89 00 00	04 17 39.0	88 55 56
NCP 5_{A2}	05 27 11.0	88 59 34	04 43 13.8	88 56 47
NCP 7_{A1}	06 32 49.0	88 59 34	05 46 28.9	89 00 04
NCP 7	07 00 00.0	89 00 00	06 13 21.4	89 01 52
NCP 7_{A2}	07 27 11.0	88 59 34	06 41 33.1	89 02 49
NCP 9_{A1}	08 32 49.0	88 59 34	07 51 44.9	89 05 58
NCP 9	09 00 00.0	89 00 00	08 21 46.8	89 07 34
NCP 9_{A2}	09 27 11.0	88 59 34	08 53 07.9	89 08 13
NCP 11_{A1}	10 32 49.0	88 59 34	10 10 49.4	89 10 13
NCP 11	11 00 00.0	89 00 00	10 43 47.2	89 11 11
NCP 11_{A2}	11 27 11.0	88 59 34	11 17 21.5	89 11 06
NCP 13_{A1}	12 32 49.0	88 59 34	12 38 38.0	89 11 07
NCP 13	13 00 00.0	89 00 00	13 12 13.2	89 11 13
NCP 13_{A2}	13 27 11.0	88 59 34	13 45 12.9	89 10 15
NCP 15_{A1}	14 32 49.0	88 59 34	15 03 00.2	89 08 17
NCP 15	15 00 00.0	89 00 00	15 34 24.0	89 07 39
NCP 15_{A2}	15 27 11.0	88 59 34	16 04 28.8	89 06 03

[a] Positions for A1 and A2 give reference beam centers (see Fig. 3).

TABLE 3

SUMMARY OF OBSERVATIONS

Session	Dates	Usable Time[a]
Jan85	1984 Nov 4–17 1984 Dec 17–1985 Feb 2	2.28
Nov85	1985 Nov 2–1985 Dec 2	4.14
Dec85	1985 Dec 12–1985 Dec 27	6.05
Dec86	1986 Dec 1–7	1.93
Jan87	1987 Jan 7–20	2.15

[a] Integration time in days on all fields used to produce the results of Table 4 in § VI. The total for all sessions is 16.6 days, corresponding to ~24 days of perfect observations with a typical observing efficiency of 70%.

a) The Water Vapor Radiometer

During the Jan85, Nov85, and Dec85 sessions a water vapor radiometer at the observatory measured atmospheric emission from sky dips through the north celestial pole about once every 15 minutes. Figure 5 shows the strong correlation between sky temperature at 20.7 GHz and data quality. The transition from good to bad data was not sharp, and typically occurred somewhere between 8 and 14 K per air mass. Although the correlation between sky temperature and fractional cloud cover was also strong, there were many clear periods during which the sky temperature was high and the data were poor. More surprising to us was the fact that good data were sometimes obtained even when the sky temperature was high. It turned out that when this occurred, the scatter in successive sky temperature measurements was always low, indicating that the atmosphere was uniformly stratified. It is clear that the water vapor radiometer provides an objective, independent check of the data quality, and increases the efficiency of the observations.

b) Preliminary Editing

Two hour (i.e., single-field) blocks of data were rejected according to four criteria, none of which can bias the results. Specifically, data were rejected if:

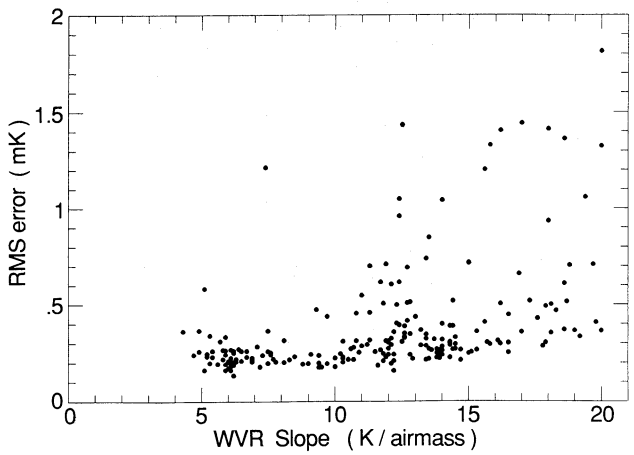

FIG. 5.—rms scatter in 1 hr blocks of data vs. atmospheric temperature per air mass measured by the water vapor radiometer at 20.7 GHz. Above 11 K per air mass the likelihood of large scatter increases significantly; however, if the error in the sky temperature measurement (not shown) was less than 0.4 K per air mass (indicating a smooth, stable atmosphere), the scatter in the data remained low. Data taken when the sky temperature and error were above 11 and 0.4 K per air mass, respectively, were not used.

1. There was an equipment failure, or any receiver adjustments were set to nonstandard values.

2. The weather was bad. If water vapor radiometer measurements were available, data were rejected if the sky temperature exceeded 11 K per air mass and the scatter in successive measurements exceeded 0.4 K per air mass. If WVR measurements were not available, data were rejected if there were clouds anywhere except low in the west over the Sierra Nevada within 2 hr of the observations. On dark nights thin clouds cannot be seen, so all data taken at night were rejected if clouds were visible at subset or sunrise.

3. The rms scatter in a 2 hr block of data indicated a noise level greater than 16 mK $s^{1/2}$. This is twice the thermal noise level under the absolute best conditions when $T_{sys} = 40$ K at the pole, and ~ 1.8 times the thermal noise level under typical very good conditions. Two hour blocks contained up to 66 successive observations of a single field. Since the rms value can be perturbed seriously by widely discrepant measurements, any measurement more than 3 times the rms distance from the mean was excluded, and the mean and scatter recalculated, to convergence. (If the scatter with rejection was below twice the thermal value, however, *all* data in the block were kept for analysis by the procedure of § IVc.) There was rarely an obvious cause of bad data rejected by this criterion. Nevertheless, such a noise level is a sure sign of trouble. We emphasize that this criterion applies to the *scatter* of the measurements in a 2 hr block of data, not to the deviation of the 2 hr mean from the mean of all data for a given field. By this criterion, therefore, we do not excise outlying points in the overall distribution for a given field, rather we eliminate *all* data taken when the scatter is large. If the excess noise is random, the only effect of this rejection will be to reduce the total integration time. If, as is much more likely, the excess noise is not random (e.g., noise from persistent structures in the atmosphere), this rejection protects the final data set against the introduction of measurements with systematic errors large compared to the final estimated error in the mean for a given field (~ 30 μK), but small compared to the width of the distribution of all measurements (9 mK $s^{1/2} \times [160\ s]^{-1/2} \approx 700$ μK). In the best weather almost no data were rejected by this criterion; however, when the water vapor radiometer was not working this criterion was our only protection against the condition described above of clear skies but large atmospheric fluctuations.

4. The number of data points in 2 hr was less than one-third the expected number, and no cause harmless to data quality could be determined.

One 40 minute period of data was rejected because the mean value was more than 3 σ away from the average of all the data for this field.

c) *Final Editing and Data Reduction*

The editing method just described easily removes data ruined by causes that persist for more than 30 minutes or so, but would be tedious to use for the removal of isolated bad points. In the best weather when the receiver appears to be working perfectly, the measured values of ΔT_i have a Gaussian distribution with width near the thermal noise limit, and the values of σ_i are all about the same. In merely good weather the distribution remains Gaussian, but the width increases consistent with the increase in the values of σ_i. The automatic procedure described below rejects individual values of ΔT_i or σ_i that are inconsistent with the underlying Gaussian distribution.

Let ΔT_i and σ_i, $i = 1, \ldots, N$ be the calibrated mean values and their standard deviations for a given field. Suppose that these N measurements were made in j 2 hr blocks. (During a complete observing day there would be two blocks each for fields 5–11, and one each for fields 1, 3, 13, and 15.) Let n_j be the number of measurements in the jth block (note that $n_j \leq 66$). Then $N = \sum_j n_j$. We calculate the weighted mean of the jth 2 hr block of data

$$B_j = \frac{\sum_{i \in J} \sigma_i^{-2} \Delta T_i}{\sum_{i \in J} \sigma_i^{-2}},$$

and a measure of the scatter about the mean

$$\sigma_j^2 = \frac{\sum_{i \in J} \sigma_i^{-2} (\Delta T_i - B_j)^2}{\sum_{i \in J} \sigma_i^{-2}},$$

where J is the set of all i in the jth block. Note that if all the σ_i are the same, and are accurate measures of the errors in the ΔT_i, then σ_j will be equal to σ_i. Systematic errors in the ΔT_i not reflected in σ_i will give $\sigma_j > \sigma_i$.

We assign a weight to each ΔT_i given by

$$w_i = (\sigma_i^2 + \sigma_j^2)^{-1}.$$

Since σ_i is a measure of the scatter in 0.5 s integrations, and σ_j is a measure of the scatter in 80 s integrations, both short- and medium-term fluctuations are represented in the weight. We calculate the weighted mean temperature

$$\overline{\Delta T} = \frac{\sum_{i=1}^{N} w_i \Delta T_i}{\sum_{i=1}^{N} w_i}, \tag{4a}$$

the estimated variance of the mean

$$\sigma^2 = \frac{1}{N-1} \frac{\sum_{i=1}^{N} w_i (\Delta T_i - \overline{\Delta T})^2}{\sum_{i=1}^{N} w_i}, \tag{4b}$$

the rms average of all σ_i

$$\sigma^* = \left(\frac{1}{N} \sum_{i=1}^{N} \sigma_i^2 \right)^{1/2},$$

and the rms average of the σ_i for each 2 hr block

$$\sigma_j^* = \left(\frac{1}{n_j}\sum_{i\in J}\sigma_i^2\right)^{1/2}.$$

The ith measurement is rejected if $|\Delta T_i - \overline{\Delta T}| \geq q\sigma_i$ or $\sigma_i \geq 2\sigma^*$, where q is a constant whose initial value is arbitrarily chosen to be 4.0. The first criterion eliminates measurements whose deviations from the mean are too large to be statistically plausible based on the scatter in the 160 constituent 0.5 s samples. The second eliminates measurements with extremely large scatter in the 160 samples. The factor of 2 in this criterion is justified by the distribution of σ_i/σ^* in Figure 6, which shows that any value near 2 would distinguish normal from discrepant values.

Similarly, the jth block of data is rejected if $|B_j - \overline{\Delta T}| \geq q\sigma_j(n_j - 1)^{-1/2}$ or $\sigma_j \geq 2\sigma_j^*$. Once again, the first criterion eliminates blocks of data with a mean implausibly far from the overall mean, based on the scatter in the n_j 80 s values. The second eliminates blocks of data whose 80 s values have large scatter in their constituent samples.

The above procedure is repeated to convergence, with deleted measurements or blocks added back in as necessary, for $q = 4.0, 3.9, \ldots, 2.7$. About 5% of the measurements are rejected when $q = 4.0$, and $\sim 10\%$ when $q = 3.0$. A common cause of individual measurements rejected by this procedure is loss of LO phase lock. *All results reported in this paper are for* $q = 3$. This choice is based on the following consideration. There are on average 2000 measurements per field in our final data set. If the σ_i were the true errors, we would expect about five measurements per field to be rejected erroneously with $q = 3$. In fact, the σ_i typically underestimate the true errors by $\sim 30\%$ (see § V), so we expect ~ 40 measurements per field to be rejected erroneously. This will cause a slight reduction in the variance of the final data set, but it cannot introduce a bias. The other 160 or so measurements per field that are rejected with $q = 3$, however, cannot be drawn from the underlying Gaussian distribution of the good data, and could easily introduce a bias.

The catalog of causes of bad data is, without elaboration: bad weather; refrigerator failure; compressor failure; compressor fan cycling; pressure variations in the refrigerator JT return line; temperature oscillations of the 4 K stage; loss of LO phase lock; and high voltage relay glitches. We believe that bad data produced by any of these have been recognized and deleted. The important question of whether there are unidentified sources of systematic error will be addressed next.

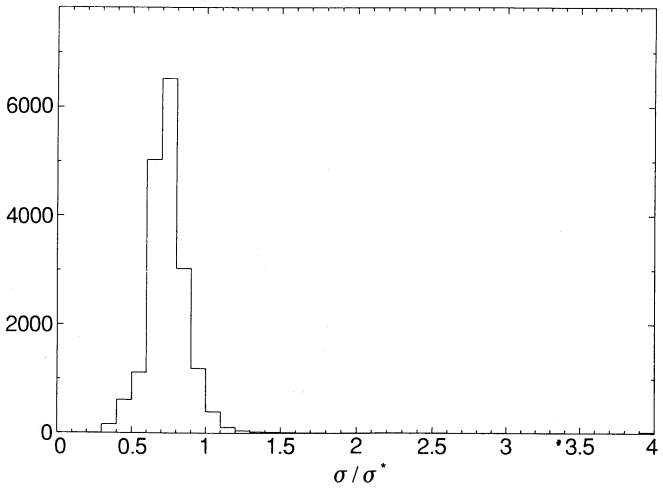

FIG. 6.—Distribution of σ_i/σ^* for all fields and all sessions. The last bin includes all values greater than 3.9. Data with $\sigma_i/\sigma^* > 2$ are deleted.

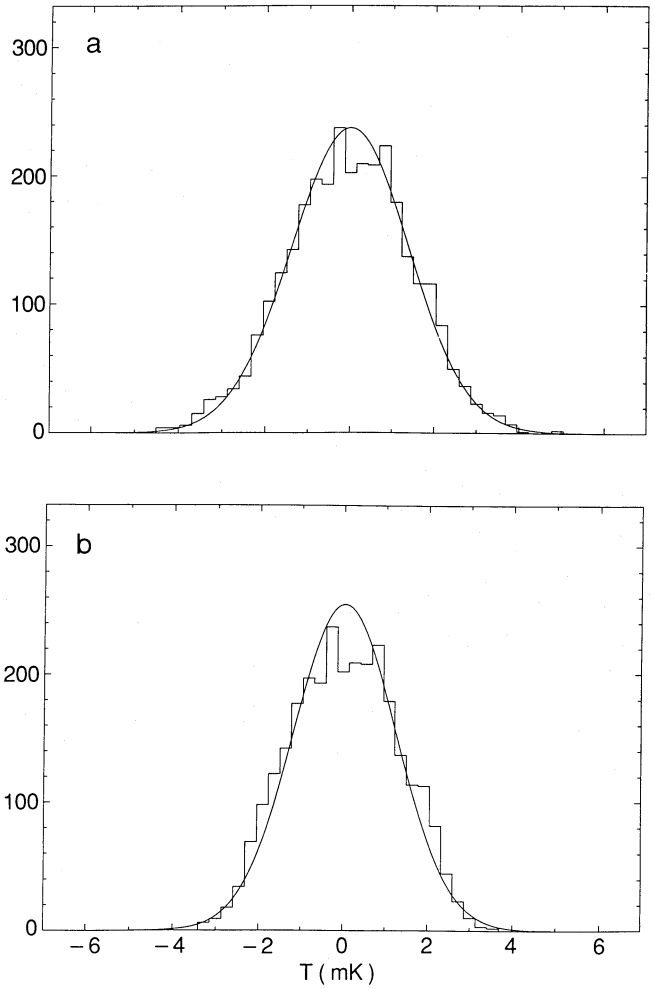

FIG. 7.—Distribution of ΔT_i from NCP 9 accepted by the procedure of § IVc with (*a*) $q = 4.0$; (*b*) $q = 3.0$. The width of the distribution, given by $\sigma_{\rm dist}^2 \equiv \Sigma w_i(\Delta T_i - \overline{\Delta T})^2/\Sigma w_i$ (eq. [4]), is 1.38 mK for $q = 4.0$ and 1.22 mK for $q = 3.0$. The curves superposed on the distributions are Gaussians with dispersion $\sigma_{\rm dist}$.

V. TESTS OF DATA QUALITY

The greatest danger in measurements such as these is persistent systematic effects with a nonzero mean. Such effects may be difficult to detect, since deviations from zero of even $T_{\rm sys} \times 10^{-6}$ are intolerable, yet only the data can reveal them. In this section we discuss the answers to four questions: (1) Are the ΔT_i normally distributed? (2) What are the time scales for systematic errors? (3) Is there a diurnal component in ΔT_i? (4) Is the mean value of all measurements for all fields zero?

As will be discussed in § VI, the observations of NCP 7 are corrupted by a weak radio source in the wings of the beam, and must be excluded from some of the tests of this section.

Figure 7 shows the distribution of accepted measurements (~ 2 days total integration time) for one of the fields for two values of q, the rejection parameter from § IVc. The plotted

curves are Gaussians with the same variance as the measurements. Specifically, $\sigma^2_{\rm dist} \equiv \Sigma w_i(\Delta T_i - \overline{\Delta T})^2/\Sigma w_i$ is 1.38 mK for $q = 4.0$, and 1.22 mK for $q = 3.0$. A measurement $\Delta T_i \pm \sigma_i$ is rejected if $|\Delta T_i - \overline{\Delta T}| \geq q\sigma_i$. (Blocks of data far from the mean compared to their internal scatter are similarly rejected.) However, this truncation does not lead to sharp edges in the distribution in Figure 7*b* because of the width of the distribution of σ_i shown in Figure 6. We find that $\sigma_{\rm dist}$ can be up to 10% larger than the dispersion for 2 hr blocks of data, and up to 50% (typically 30%) larger than the dispersion for 80 s of data. From this it is clear that fluctuations on time scales longer than 80 s have a significant effect on the width of the distribution, and therefore that σ_i underestimates the true errors. The fair agreement between $\sigma_{\rm dist}$ and the dispersion for 2 hr blocks of data shows that there are no dominating fluctuations on time scales longer than 2 hr, a conclusion supported also by Figure 8. The distributions are sufficiently Gaussian that the procedure of § IV*c* is reasonable. The question of systematic biases from long-term fluctuations must be answered by other tests given blow.

Figure 8 shows the relationship between integration time t and variance σ^2. The value of $t\sigma^2$ rises from the typical thermal limit on time scales less than a few minutes to $(1.2)^2$ times this limit on a time scale of tens of minutes, and then remains at this level for integration times up to 17 days. This shows that there are systematic errors, which we attribute to the atmosphere, with characteristic time scales of minutes to hours, but that no longer term systematic effects are seen in our data.

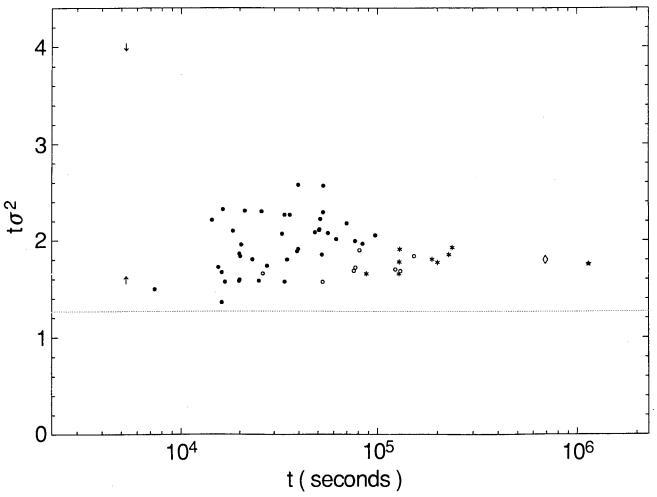

FIG. 8.—Test for correlated errors. For independent data points with uncorrelated errors, we expect $\sigma^2 \propto t^{-1}$, where σ^2 is the variance of the mean of a series of observations extending over a total integration time t. We have plotted $t\sigma^2$ for (1) each field for each observing session (●); (2) the concatenation of all sessions for each field (∗); (3) the concatenation of all sessions except Dec85 for each field (○); and (4) the concatenation of all fields for all sessions (★), and for all sessions except Dec85 (◇). Also shown is $t\sigma^2$ for one 2 hr block of data in very good weather (↑). The plot is normalized to the thermal noise level for the best $T_{\rm sys}$ ever achieved during the observations (i.e., 40 K at the pole), and the horizontal line shows the thermal noise level for more typical performance of $T_{\rm sys} = 45$ K at the pole. Although the measurements in the concatenated data sets are the same as for individual fields and sessions, there could have been systematic differences in mean values for different sessions and fields, giving larger values of σ for the concatenated data sets. That is not the case. Two hour blocks of data with variances exceeding four times the 40 K thermal level (↓) are edited out (see § IV*b*). Correlated atmospheric fluctuations on time scales of minutes to 2 hr increase the noise to ~20% above the typical thermal limit. There is no sign of correlated errors for times longer than 2 hr.

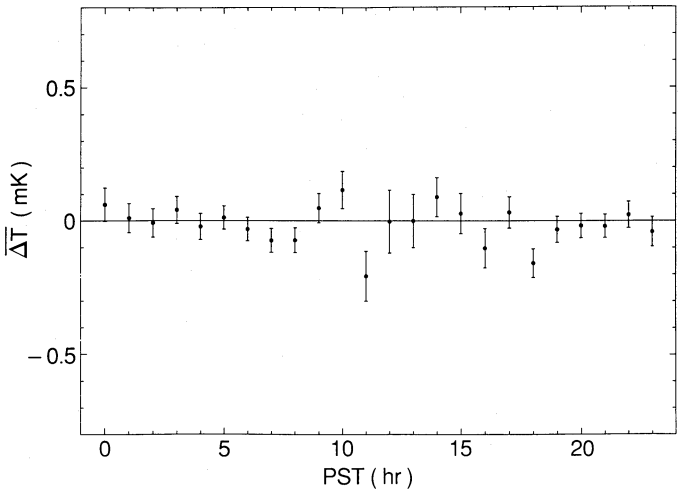

FIG. 9.—$\overline{\Delta T} \pm \sigma$ as a function of time of day. Data accepted by the procedures of § IV (with $q = 3.0$) for all fields except NCP 7 have been combined, then divided into 1 hr blocks according to the PST of the observations (0^h–1^h, etc.). The weighted mean and error for each block were found using the procedure of § IV*c*. The size of the error bars varies with the number of measurements in the block. There are fewer daytime measurements, because the daytime atmosphere is noisier, and because instrumental adjustments are scheduled during the day. For example, bins 4–6 contain 2491 measurements, while bins 11–13 contain only 587. For 24 degrees of freedom $\chi^2_\nu = 0.83$, and there is little evidence for diurnal systematic errors.

Figure 9 shows the mean temperature and error as a function of time of day. This is the best way to detect systematic errors caused by the Sun. No systematic deviations from zero are seen, although the scatter in the measurements is larger during the day than at night, and, as a result, more measurements have been rejected by our editing procedures.

Figure 10 shows the mean temperature and error as a function of observing session. This is the best way to detect low-level systematic errors that persist for weeks. The expected

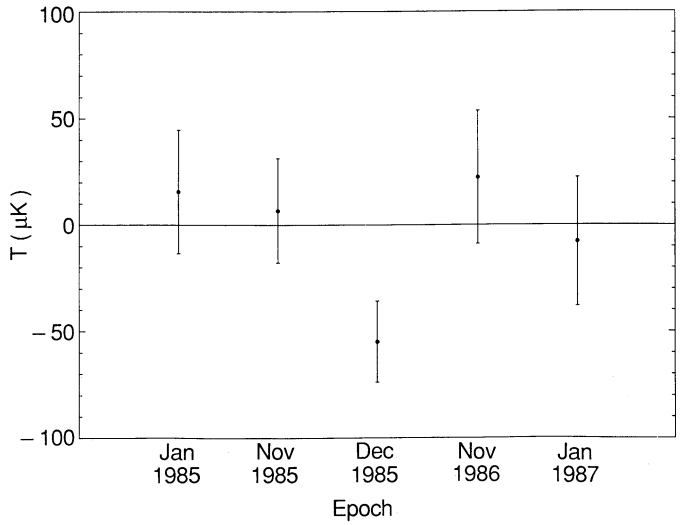

FIG. 10.—$\overline{\Delta T} \pm \sigma$ for the five observing sessions. For each session, data accepted by the procedures of § IV (with $q = 3.0$) for all fields except NCP 7 have been combined, and the weighted mean and error found using the procedure of § IV*c*. The reduced χ^2 value is 1.85 for 5 degrees of freedom, and the probability of a value for χ^2 greater than this is 10% for a Gaussian distribution.

mean for many fields in differential measurements is zero. Primarily because of the low mean for Dec85, the probability of such a set of five values from Gaussian distributions is 10%. Despite a concerted effort to uncover sources of systematic error that could produce offsets of tens of microkelvins, none has been found. The Dec85 mean value is not so low as to be manifestly due to bad data, yet it remains the most serious strain in our claim that no significant systematic errors bias the data. The results from several more observing sessions should clarify this situation. When run on the combined data from all fields except NCP 7, the procedure of § IV*c* gives $\overline{\Delta T} = -11 \pm 10$ μK, not inconsistent with zero. It should be noted that an instrumental bias is unlikely to cancel a real temperature difference in the microwave background radiation, so that the upper limits derived below are likely to overestimate the limit on the true sky noise.

VI. RESULTS

Table 4 and Figure 11 give the results of the OVRO measurements. These are thermodynamic temperature differences on the sky, including a 4% correction for atmospheric absorption and a 1% correction for the difference between $\partial I_{\text{Rayleigh–Jeans}}/\partial T$ and $\partial I_{\text{Planck}}/\partial T$ at $T = 2.78$ K. The most obvious feature is the 8 σ mean in field NCP 7. No measurement of any other field in any session has such a significant nonzero value. We believe that the only reasonable interpretation is that there is a real temperature difference on the sky near this field.

TABLE 4

OVRO MEASUREMENTS

Field	$\overline{\Delta T} \pm \sigma$ (μK)
NCP 1	-64 ± 35
NCP 3	20 ± 34
NCP 5	-29 ± 27
NCP 7	217 ± 28
NCP 9	34 ± 26
NCP 11	-23 ± 26
NCP 13	-20 ± 32
NCP 15	-36 ± 39

VII. EXTRANEOUS SOURCES OF ANISOTROPY

We have discussed at some length the procedures used to control systematic errors arising from receiver or atmospheric noise and believe that no such errors remain in the edited data at a level of a few tens of microkelvins. Discrete radio sources, the Sunyaev-Zel'dovich decrement due to hot gas in galaxy clusters, galactic synchrotron emission, and interstellar dust are possible sources of anisotropy contributing to the measurements in Table 4. For even a modest number of fields, discrete sources or clusters are unlikely to cancel fluctuations in the microwave background. Thus anisotropy estimates based on data uncorrected for such sources are likely to overestimate the anisotropy of the microwave background itself.

The total anisotropy produced by clusters through the Sunyaev-Zel'dovich effect depends on the redshift at which

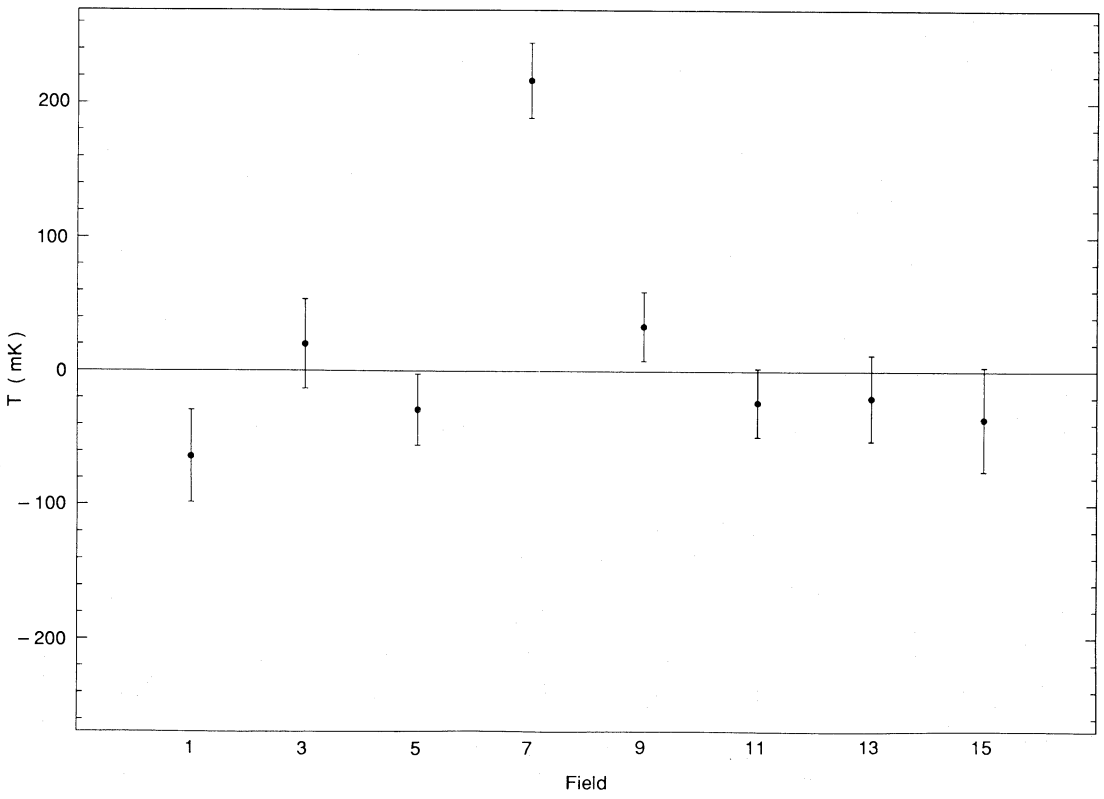

FIG. 11.—Results of the OVRO measurements for fields NCP 1–NCP 15 after data editing as described in § IV. The result for NCP 7 deviates significantly from zero, because of a confusing source (PT 58, see § VII). The remaining seven fields are consistent with zero.

clusters formed. Since this is not known, no *a priori* estimates can be made. Instead, limits on the anisotropy of the microwave background provide useful constraints on the redshift at which clusters formed (Sunyaev 1977, 1978; Rephaeli 1981).

The effects of discrete sources could be subtracted, using the beam map in Figure 4*a*, if their flux densities *during the microwave background observations* were known; however, for weak sources accurate flux densities require long integrations. Given that at 20 GHz we expect many sources to be variable, such long integrations would have to be repeated at every observing epoch, and flux density observations could easily dominate the observing program. Moreover, at flux levels where the density of sources on the sky is greater than about one source per 10 beam areas, a different telescope with higher resolution must be used. Based on source-count estimates of Danese, De Zotti, and Mandolesi (1983) and Franceschini *et al.* (1989) the sensitivity limit for the 40 m telescope set by discrete sources will be in the range 15–45 μK. It remains to be seen whether the limit set by clusters is even higher.

The 8 σ detection in NCP 7 is far from the confusion limit and is well above previous upper limits on microwave background anisotropy. It is due to source 58 in the 4.85 GHz Pauliny-Toth *et al.* (1978) survey of the north celestial pole. Table 5 gives all sources in this survey within 10′ of any field in Table 2, along with 20 GHz flux densities measured at OVRO. In the beam map in Figure 4*a* PT 58 appears at the 3% level for upper culmination observations and the 6% level for lower culmination observations. Thus for the range of measured flux densities in Table 5, PT 58 should produce temperatures in NCP 7 of 130–380 μK—just what has been found. Indeed, it became clear early in the course of these observations that there was a variable source in field NCP 7; however, since a significant positive result provided reassurance that the system was working correctly, we continued to observe the field. Therefore NCP 7, the only field that was suspect on the basis of these observations alone, has been excluded from the analysis that follows.

The only other source in the Pauliny-Toth sample that affects our data is PT 82, 1.4 from the center of NCP 9_{A2}. From the 20 GHz flux densities in Table 5 and the beam map we would expect a -28 to -49 μK shift in NCP 9. The results on this field show no such effect. However, when the data for NCP 9 are divided according to parallactic angle, we measure -53 ± 36 μK at parallactic angles for which the reference beam covers PT 82, and 78 ± 30 μK at angles for which PT 82 does not affect the measurements. From this we deduced that a source too weak to appear in the Pauliny-Toth catalog lay near the main field of NCP 9. A map of this region at 5 GHz, made for us by P. Crane with the VLA, does indeed show a small-separation double $\sim 75''$ from NCP 9. A source at this level with a spectral index of -0.75 ± 0.15 (see, e.g., Donnelly, Partridge, and Windhorst 1987) would raise the temperature of NCP 9 by 90 ± 20 μK. Our direct 20 GHz measurements of PT 82 show that it is variable, so we cannot correct the NCP 9 data for it. If we take the 78 ± 30 μK value for parallactic angles at which PT 82 does not affect the data, and apply a 90 ± 20 μK correction for the double (which is unlikely to be a variable source), we obtain -12 ± 36 μK for NCP 9.

While the Pauliny-Toth *et al.* survey has helped to focus our attention on the problems of discrete sources, correcting for such sources requires observations down to 1 mJy, much below the Pauliny-Toth *et al.* cutoff at 14 mJy. At present such correction, based on the VLA image, is possible only for NCP 9. Using -12 ± 36 μK in place of the value in Table 4 gives a limit $\sim 6\%$ lower than that derived in § VIII. If NCP 9 were excluded entirely from our analysis, our derived upper limit on sky fluctuations would increase by 10%, because of the decrease from seven to six fields. Thus the choice between making no corrections to NCP 9, correcting NCP 9, or excluding NCP 9 entirely has no important consequence. We have chosen the first course, since in doing this we reject only the one field (NCP 7) for which there is clear evidence *in our own observations* of a variable discrete source, and since we prefer to delay the correction of NCP 9 until we can treat all fields equally using VLA maps of comparable sensitivity. Accordingly, in the analysis below we use the seven values in Table 4 excluding NCP 7.

The successful prediction and subsequent detection of a source in the main beam of NCP 9 gives us confidence that our data can be trusted at the 50 μK level. In the future it will be possible to subtract the effect of sources like the double near NCP 9, which are unlikely to vary; however, variable sources like PT 58 and PT 82 must be avoided.

VIII. ANALYSIS

Table 4, along with Figures 2 and 4 and equation (4), provides a complete summary of our results. The question that we now turn to is: what is the largest anisotropy of the microwave background radiation consistent (in some sense to be specified) with the seven measurements in Table 4 (excluding NCP 7). This is a statistical question, and its answer, and the most appropriate method for deriving it, depends on the unknown distribution of sky fluctuations. We will assume Gaussian fluctuations because this provides a common starting point for comparison of our results with those of other observers and with model predictions, which, with few exception (e.g., cosmic

TABLE 5

SOURCES FROM THE PAULINY-TOTH[a] 4.85 GHz SURVEY WITHIN 10′ OF NCP FIELDS

SOURCE	EPOCH 1985.0 A.	EPOCH 1985.0 Decl.	EPOCH 1950.0 R.A.	EPOCH 1950.0 Decl.	FLUX DENSITY (mJy) 4.85 GHz[a]	FLUX DENSITY (mJy) 20.0 GHz[b]	RELATIVE POSITION
PT 42.....	$05^h38^m14^s.8$	89°04′51″	$04^h49^m46^s.0$	89°02′30″	28.8	9–15	6.0 from NCP 5_{A2}
PT 58.....	06 56 42.8	88 58 14	06 11 21.0	88 59 58	26.0	18–25	2.0 from NCP 7
PT 82.....	09 28 26.3	88 58 06	08 55 28.0	89 06 48	18.2	1.4–2.4	1.4 from NCP 9_{A2}
PT 87.....	10 51 21.1	89 03 31	10 31 56.0	89 14 32	71.5	12–15	4.1 from NCP 11
PT 90.....	14 19 54.6	88 52 22	14 44 19.0	89 01 35	18.8	6	8.1 from NCP 15_{A1}
PT 91.....	15 01 11.5	89 08 32	15 42 46.0	89 16 00	28.4	5	8.5 from NCP 15

[a] Pauliny-Toth *et al.* 1978. The flux density limit was 14 mJy; position errors were typically 15″.
[b] From OVRO, 1987 Jun, 1988 Jan, and 1988 Feb. Typical errors are ± 0.4 mJy.

strings) assume Gaussian fluctuations. Moreover, as we show in the Appendix observations of seven fields are well-suited to the detection of Gaussian fluctuations, but any power-law tail in the fluctuation spectrum with an index greater than -2 would be detected with greater probability in an experiment with more fields.

We first derive a limit on $\theta^2_{\rm sky}$, the variance of the distribution of sky fluctuations from the mean background temperature *for the triple beam* of Figure 4. We then calculate corresponding limits on the autocorrelation function of the microwave background fluctuations, under simple assumptions.

a) Limits on $\theta^2_{\rm sky}$

Arguments about the "best" statistical procedure to use in deriving limits on $\theta^2_{\rm sky}$ can be misleading, because statistical procedures that are optimal according to some criterion may be inappropriate if the assumptions on which they are based are not justified. We use two different procedures to derive limits on $\theta^2_{\rm sky}$. One (the likelihood ratio test) has been used widely in past microwave background work, the other (likelihood) has not. The limits given by the two procedures are about the same. However, as we discuss at the end of this section, the relative insensitivity of the likelihood procedure to certain problems that may afflict real data sets makes it the preferred procedure for microwave background work.

i) *Likelihood*

The variance of the (assumed Gaussian) distribution of sky fluctuations for the triple beam in Figure 4 is $\theta^2_{\rm sky}$. The mean of the distribution of fluctuations is zero, since the triple beam samples only deviations from mean sky temperature. Thus the probability density for a measurement $\Delta T \pm \sigma$ is

$$p(\Delta T, \sigma) = [2\pi(\sigma^2 + \theta^2_{\rm sky})]^{-1/2} \exp\left[-\Delta T^2/2(\sigma^2 + \theta^2_{\rm sky})\right] .$$

The joint density for the seven fields (excluding NCP 7) is

$$L(\{\Delta T_i\} \mid \theta_{\rm sky}) \equiv \prod_{i=1}^{7} p(\Delta T_i, \sigma_i)$$

$$= \prod_{i=1}^{7} [2\pi(\sigma_i^2 + \theta^2_{\rm sky})]^{-1/2} \exp\left[\frac{-\Delta T_i^2}{2(\sigma_i^2 + \theta^2_{\rm sky})}\right] .$$

$L(\{\Delta T_i\} \mid \theta_{\rm sky})$, called the *likelihood function*, can be thought of as the relative probability of the set of seven measurements as a function of the assumed sky variance. This function, normalized to its maximum value, is given in Figure 12 for the measurements in Table 4.

The value of $\theta_{\rm sky}$ for which our results have the maximum likelihood is 14 μK; however, $L(\theta_{\rm sky} = 0)$ is almost as large, so that this cannot be claimed as a detection. Upper limits can be determined in two ways. The first is to find the value of $\theta_{\rm sky}$ at which the relative likelihood has fallen to some specified value, say e^{-2}, 0.05, or 0.01. From Figure 12, these relative likelihood values occur at 57, 72, and 98 μK, respectively.

Alternatively, we can use Bayes's formula, which gives the probability density of $\theta_{\rm sky}$ given the set of observations $\{\Delta T_i\}$ as

$$p(\theta_{\rm sky} \mid \{\Delta T_i\}) \propto L(\{\Delta T_i\} \mid \theta_{\rm sky}) p(\theta_{\rm sky}) ,$$

where $p(\theta_{\rm sky})$ and $p(\theta_{\rm sky} \mid \Delta T_i)$ are known, respectively, as the *prior* and *posterior* densities (see Berger 1985 and Berger and Wolpert 1984 for general discussions of Bayesian methods; and Edwards 1984 for an alternative interpretation of the likelihood function). Our (possibly nonexistent) knowledge of $\theta_{\rm sky}$ before the observations is represented by $p(\theta_{\rm sky})$. The simplest prior density consistent with the one fact we know (i.e., $\theta^2_{\rm sky} \geq 0$), is $p(\theta_{\rm sky}) = c$ for $\theta_{\rm sky} \geq 0$, and $p(\theta_{\rm sky}) = 0$ for $\theta_{\rm sky} < 0$. Then the posterior density, i.e., the density of $\theta_{\rm sky}$ determined from what we knew before the observations plus the observations themselves, will look just like Figure 12 except for normalization. In this case, the area under the curve from 0 to θ^* gives the probability that $0 \leq \theta_{\rm sky} \leq \theta^*$. Specifically, $P(\theta_{\rm sky} < 58$ μK$) = 0.95$, and $P(\theta_{\rm sky} < 127\ \mu$K$) = 0.9987$. Thus 58 μK can be taken as a 95% upper limit to $\theta_{\rm sky}$, and 127 μK can be taken as an equivalent 3 σ upper limit.

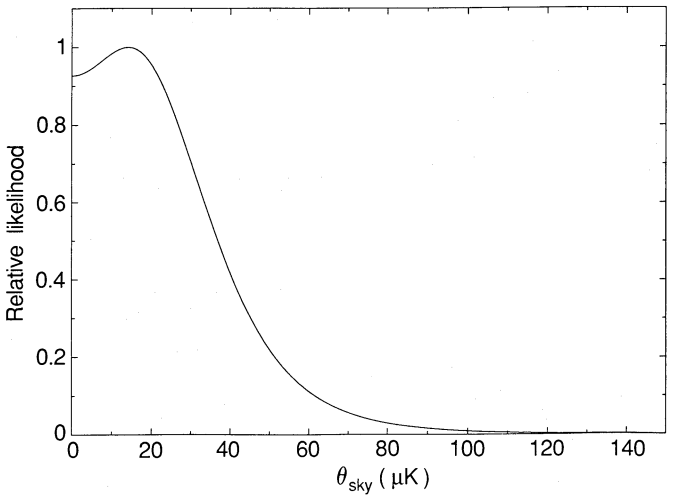

FIG. 12.—Likelihood function for the measurements in Table 4 excluding NCP 7, normalized to the maximum value. The maximum occurs at $\theta_{\rm sky} = 14$ μK, but $L(0)$ is so large that no detection of fluctuations can be claimed.

Invariance arguments suggest that when no other information is available the prior density of a parameter that is a *scale factor* (as is $\theta_{\rm sky}$ in the absence of measurement errors, see Berger 1985) should be uniform in θ^{-1} rather than θ. With such a prior density, a cutoff must be imposed at small values of $\theta_{\rm sky}$ to make the posterior density normalizable, and the derived upper limit to $\theta_{\rm sky}$ depends somewhat on the value chosen. This is unappealing, but it is not a serious problem if observations of galaxies and clusters do provide a physical basis for such a cutoff. We find 95% limits in the range 30 μK to 36 μK for cutoffs between 0.1 μK and 1.0 μK. However, the measurement errors fix a level below which the actual value of $\theta_{\rm sky}$ makes little difference, and what we measure for each field is determined almost entirely by the errors. Yet a $1/\theta$ prior distribution gives much more weight to values of $\theta_{\rm sky}$ below this level than it does to large values where our measurements have real discriminating power, leading to quite low limits on $\theta_{\rm sky}$. We therefore prefer a conservative prior density uniform in $\theta_{\rm sky}$, and will adopt the corresponding limit of 58 μK.

Figure 12 shows $L(\{\Delta T_i\} \mid \theta_{\rm sky})$ only for $\theta^2_{\rm sky} \geq 0$. Whether we think of this as a restriction of the prior density or the likelihood function itself to positive sky variances makes little difference. The ease with which this physical constraint can be imposed makes the likelihood function estimate relatively insensitive to a problem that can seriously affect the method described in the next section.

ii) *Likelihood Ratio Tests*

Likelihood ratio tests have been used for analysis of microwave background data by many authors, e.g., Boynton and

Partridge (1973), Lasenby and Davies (1983), and Uson and Wilkinson (1984*a*, *b*, *c*). For a general discussion of likelihood ratio tests and definitions of the statistical terms used here, see Lehmann (1986).

The problem is to choose between an hypothesis H and an alternative hypothesis K. Specifically, given a set of measurements whose distribution is known in terms of a parameter θ, we must find θ. We assume that if θ is known, it is known whether H is true. A "test" is a set of criteria used to accept or reject H. The *level of significance* or *size* α of the test is the probability of rejecting H if H is true, i.e., a type I error. The *power* β of the test is the probability of rejecting H if K is true, so that $1 - \beta$ is the probability of a type II error (i.e., of accepting H when K is true). Traditionally it has been assumed that type I errors are worse than type II errors, and tests have been devised that maximize β for a fixed value of α. Such tests are called *most powerful.*

For a random sample x drawn from a distribution characterized by a parameter θ, the Neyman-Pearson lemma shows that there exists a most powerful test of size α of the simple hypothesis $H:\theta = \theta_0$ against the simple alternative $K:\theta = \theta_1$, and that this test is given by the prescription:

$$\text{reject } H \text{ if } \quad \lambda^* \equiv \frac{P(x \mid H)}{P(x \mid K)} \leq k^* ,$$

or

$$\text{accept } H \text{ if } \quad \lambda^* > k^* ,$$

where k^* is given implicitly by the requirement that $P(\lambda^* \leq k^* \mid H) = \alpha$.

In general, whether λ^* is greater than k^* or not depends on θ_1, If it does not, that is, if the ordering of points in the measurement space according to their values of λ^* does not depend on θ_1, then the test is said to be *uniformly most powerful* (UMP) and can decide between composite hypotheses, e.g., $H:\theta \geq \theta_0$ and $K:\theta < \theta_0$. (Note that the test is the same as before, but the distributions satisfy additional requirements, and stronger theorems apply. We use the generic expression "likelihood ratio tests" rather than "Neyman-Pearson tests" to emphasize this distinction.)

Given our assumption of Gaussian sky fluctuations, we can calculate λ^* for the measured $\Delta T_i \pm \sigma_i$:

$$\lambda^*(\Delta T) = \frac{\prod_{i=1}^{7} p(\Delta T_i \mid \sigma_i, \theta_{\rm sky} = \theta_0)}{\prod_{i=1}^{7} p(\Delta T_i \mid \sigma_i, \theta_{\rm sky} = \theta_1)}$$
$$= \prod_{i=1}^{7} \left(\frac{\sigma_i^2 + \theta_1^2}{\sigma_i^2 + \theta_0^2}\right)^{1/2} \exp\left\{\frac{\Delta T_i^2}{2}\left[\frac{\theta_0^2 - \theta_1^2}{(\sigma_i^2 + \theta_0^2)(\sigma_i^2 + \theta_1^2)}\right]\right\} .$$

To find k^* we must know the distribution of λ^* under H. It is simpler, and involves no loss of information about θ, to calculate the distribution under H of

$$\lambda = \sum_i \frac{x_i^2}{(\sigma_i^2 + \theta_0^2)(\sigma_i^2 + \theta_1^2)} , \tag{5}$$

where x_i is Gaussian-distributed with variance $\sigma_i^2 + \theta_0^2$, and the sign has been chosen for the case $\theta_0 > \theta_1$. That is, λ is *sufficient for* θ. The observed value of λ is given by

$$\lambda(\Delta T) = \sum_i \frac{\Delta T_i^2}{(\sigma_i^2 + \theta_0^2)(\sigma_i^2 + \theta_1^2)} .$$

Then the test of size α is as follows: reject H if $\lambda(\Delta T) \leq k$, and accept H if $\lambda(\Delta T) > k$, where k is found from $P(\lambda \leq k \mid H) = \alpha$. To calculate the power β of the test, we calculate the distribution of λ under $K:\theta = \theta_1$, where this time x_i is Gaussian distributed with variance $\sigma_i^2 + \theta_1^2$. Then $\beta(\theta_1) = P[\lambda \leq k \mid \theta = \theta_1]$.

We want to use the test to find an upper limit to $\theta_{\rm sky}$ at a certain value of α. To do this, we fix α, then vary θ_0 until $P[\lambda \leq \lambda(\Delta T)] = \alpha$. In other words, we find θ_0 such that if we sample randomly from seven Gaussian distributions with variance $\sigma_i^2 + \theta_0^2$, the probability of finding λ no larger than we have measured it $[\lambda(\Delta T)]$ is only α.

It turns out that for the distribution in equation (5), the question of whether $\lambda(\Delta T) \leq k$ or not is independent of the value of θ_1. Therefore the likelihood ratio test specified above is UMP.

Figure 13 shows the results of these calculations for the data of Table 4. On this basis, we reject $H:\theta_{\rm sky} \geq 52.5$ μK at the 95% confidence level (i.e., $\alpha = 0.05$), and reject $H:\theta_{\rm sky} \geq 107$ μK at the 99.87% confidence level. The power of the test β is 0.72 (independent of α) for $\theta_1 = 0$, and decreases monotonically until $\beta = \alpha$ at $\theta_1 = \theta_0$.

Under appropriate assumptions, likelihood ratio tests are "most powerful." This means only that the power is greater than that of any other test at the same level of significance; it does *not* mean that the power is close to unity. Results of likelihood ratio tests should always be regarded with suspicion unless the power is stated along with the level of significance.

The inclusion of measurement errors in the likelihood ratio test has important consequences. While these have been pointed out by others, we restate them here for emphasis. Consider the following. The term $\sigma_i^2 + \theta_0^2$ in equation (5) comes from the combination of the measurement errors with the sky fluctuations, so that, physically, $\sigma_i^2 + \theta_0^2 \geq \sigma_i^2$. Formally, however, the term can be treated as a single quantity that satisfies only $\sigma_i^2 + \theta_0^2 \geq 0$. In effect, this "allows" sky fluctuations to compensate for measurement errors, reducing the width of the overall distribution. The likelihood ratio test requires $P(\lambda \leq \lambda(\Delta T) \mid H) = \alpha$. Suppose that the measurements

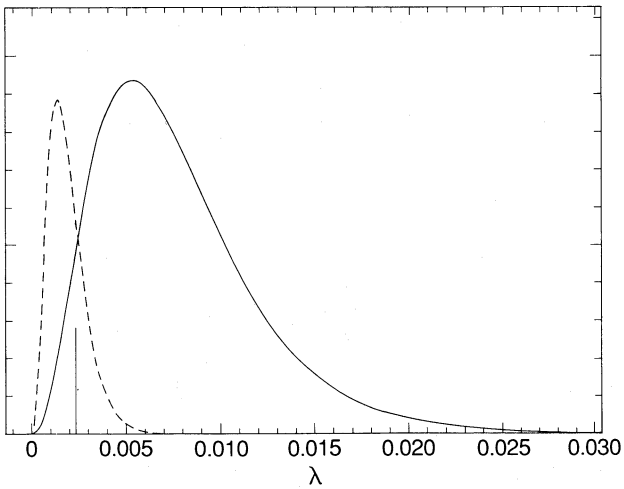

FIG. 13.—Distribution of $\lambda = \sum_i x_i^2/(\sigma_i^2 + \theta_0^2)(\sigma_i^2 + \theta_1^2)$ with $\theta_0 = 52.48$, $\theta_1 = 0$ from a Monte Carlo simulation. Each x_i was assumed to be Gaussian-distributed with variance $\sigma_i^2 + \theta_0^2$ (*solid line*) or $\sigma_i^2 + \theta_1^2$ (*dashed line*). In each case, 300,000 sets of x_i were generated. The curves are independently normalized. A vertical line marks $\lambda(\Delta T) = 0.00233$. The area under the solid curve to the left of this line is $\alpha = 0.05$, while the area to the left under dashed curve is $\beta(\theta_1 = 0) = 0.72$.

are much closer to zero than expected for the size of the measurement errors *alone* (i.e., $\chi_\nu^2 \ll 1$), and that $P[\lambda \leq \lambda(\Delta T) | \theta_0 = 0] < \alpha$. Even with $\theta_0 = 0$, the requirements of a test of size α cannot be satisfied. If, however, $\sigma_i^2 + \theta_0^2$ is treated as a single quantity that can be reduced to arbitrarily small values, a test of size α always exists (see Partridge 1980*b* for an example), even though it is physically absurd. By imposing the condition $\theta_0^2 \geq 0$ in the algorithm that performs the test, we can avoid this absurdity. But in less extreme cases the test will still implicitly treat $\sigma_i^2 + \theta_0^2$ as a single quantity and will underestimate θ_0 to compensate for scatter in the measurements that is unexpectedly small given the errors alone (see Lawrence, Readhead, and Myers 1988 for an example). Fortunately, the power of the test against the alternative $\theta_1 = 0$ immediately shows the existence of this problem. Consideration of how Figure 13 would look for various θ_0 and θ_1 shows that as θ_0 goes to zero, $\beta_{\max} = \beta(\theta_1 = 0)$ goes to α. When the power is low, a high confidence level is meaningless.

iii) *Comparison of Statistical Tests*

The OVRO data do not have the problem of $\chi_\nu^2 \ll 1$, and we therefore expect (and obtain) results from the two statistical methods that are in close agreement. It is instructive, however, to compare the performance of the methods on data sets with $\chi_\nu^2 \approx 1$, as may be obtained in practice either by chance or through misestimation of measurement errors. Such a comparison can be made at small computing cost for the idealized case in which the measurement errors for all fields are equal. Then the likelihood function is simple, and λ (see eq. [5]) has a χ^2 distribution.

Let $\sigma_{\rm true}$ and $\sigma_{\rm assumed}$ be the true (in practice unknown) and assumed measurement errors, respectively, for N fields. Suppose that $\sum_{i=1}^N \Delta T_i^2/\sigma_{\rm true}^2 = N$ for the measurements, and that $\theta_{\rm sky} = 0$. Consider first the case where $\chi_\nu^2 \neq 1$ because the errors are misestimated. Figure 14 and 15 show the upper limits produced by the two methods for $\frac{1}{3}\sigma_{\rm true} \leq \sigma_{\rm assumed} \leq 3\sigma_{\rm true}$ (corresponding to $9 \geq \chi_\nu^2 \geq \frac{1}{9}$). As the assumed error decreases, the 95% upper limit (as previously defined) given by

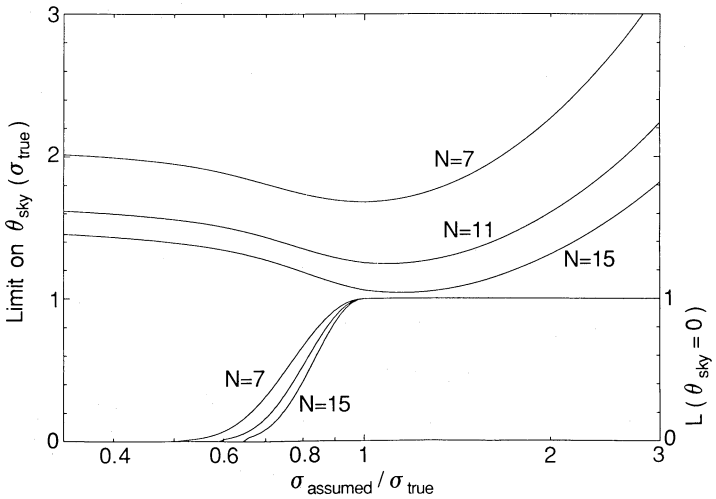

FIG. 14.—Upper limits on $\theta_{\rm sky}$ at the 95% level that would be given by a Bayesian analysis with uniform prior distribution for measurements of N fields, assuming $\theta_{\rm sky}^2 = 0$, and $\sum_{i=1}^N \Delta T_i^2/\sigma_{\rm true}^2 = N$. The abscissa is the ratio of the assumed measurement error $\sigma_{\rm assumed}$ to be the true measurement error $\sigma_{\rm true}$. The ordinate (*left-hand scale*) is in units of $\sigma_{\rm true}$. The lower family of curves gives the value of the likelihood function at $\theta_{\rm sky} = 0$ (*right-hand ordinate*). Values of $L(0)$ less than about 0.1 suggest (false) detection of fluctuations.

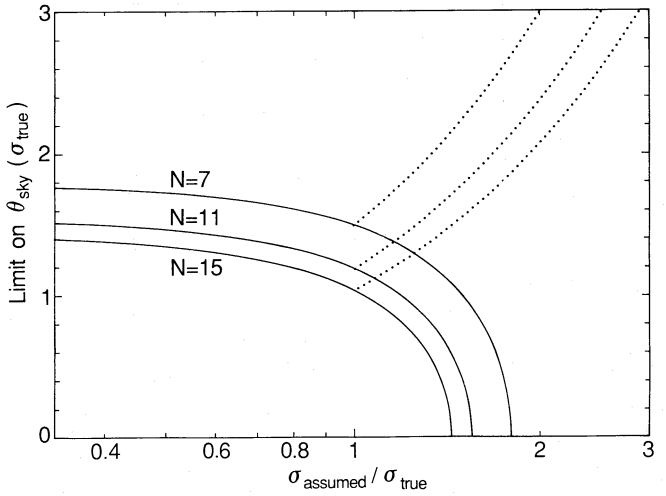

FIG. 15.—Same as Fig. 14, but for a likelihood ratio test. Dotted curves show limits given by a modified likelihood ratio test in which the power of the test is specified in advance, as described in the text. In this case the power for each N is set to the power of the "normal" likelihood ratio test for $\chi_\nu^2 = 1$. For $N = 7$, $\beta = 0.57$, while for $N = 15$, $\beta = 0.55$.

both tests increases slowly. This is desirable; however, as shown by the lower set of curves in Figure 14, with a 30% to 40% underestimate in errors the value of the likelihood function at $\theta_{\rm sky} = 0$ is only ~ 0.1, and for even lower values of $\sigma_{\rm assumed}$ *both tests would support claims of detections, rather than upper limits.* Underestimation of errors is thus a serious problem, but one to which both statistical methods are equally susceptible.

As the assumed error increases, on the other hand, the two tests behave quite differently. The 95% limit given by the Bayesian method increases, while the limit given by the likelihood ratio test *decreases to zero.* For $\sigma_{\rm assumed}$ large enough, the requirements of the likelihood ratio test cannot be satisfied for $\theta_{\rm sky}^2 > 0$. This dramatic difference between the two methods arises because the likelihood ratio test compares the measured values to what "should" have been measured given the errors, and ascribes any discrepancy to sky fluctuations, while the Bayesian method compares only changes in likelihood as $\theta_{\rm sky}$ varies. The likelihood ratio method is in some sense absolute, while the Bayesian method is relative.

The behavior of both statistical methods when data values (rather than errors) are arbitrarily changed (corresponding to the chance occurrence of larger or smaller measurements than expected from the parent distribution) follows at once, since a change of the data values alone can be thought of as a change of both data values and errors by a given factor, which changes the limits by the same factor, followed by a change of the errors while leaving the data values fixed. For example, with $N = 7$ the limits given by the Bayesian method increase by a factor of 1.12 when errors alone are multiplied by 1.5. When means are divided by 1.5, therefore, the limit will change by 1.12/1.5; i.e., the Bayesian method leads to underestimation by a factor of 1.3. For the likelihood ratio method, the limit changes by 0.66 when the errors change by 1.5, so when means are divided by 1.5 the limit will change by 0.66/1.5, i.e., the likelihood ratio method leads to underestimation by a factor of 2.3.

To summarize, underestimation of errors, or measured values by chance larger than expected given the true parent distribution, leads to higher limits or spurious detections with both statistical methods. Overestimation of errors leads to

higher limits with the Bayesian method, but lower limits or failure for the likelihood ratio test. Measured values by chance smaller than expected for the parent distribution lead to moderately lower limits for the Bayesian method, but dramatically lower limits for the likelihood ratio method. Neither method is able to extract "truth" from faulty or statistically unlikely data sets, but the results of the Bayesian method for data sets with low values of χ_ν^2 are much less misleading than those of the likelihood ratio method.

As we noted in § VIII*a*(ii), when $\chi_\nu^2 \ll 1$ the distributions of λ under $\theta_{\rm sky} = \theta_0$ and $\theta_{\rm sky} = 0$ are not very different, since both are dominated by the measurement errors, and the power of the likelihood ratio test will be low. This problem can be avoided by using the test in a different way (G. Bernstein and D. Cottingham, private communication). Rather than finding the value of θ_0 for which $\lambda_{\rm observed}$ is equal to λ at the α point of the distribution (see Fig. 13), one could find θ_0 so that the α point in the distribution of λ for $\theta_{\rm sky} = \theta_0$ coincides with the β point in the distribution of λ for $\theta_{\rm sky} = 0$. Call this value of λ $\lambda_{\alpha\beta}$. In effect, the choice of β fixes the minimum value of $\theta_{\rm sky}$ that we believe can be distinguished from zero, given the measurement errors. After specifying α and β in advance, we compare $\lambda_{\rm observed}$ with $\lambda_{\alpha\beta}$: if $\lambda_{\rm observed} \leq \lambda_{\alpha\beta}$ we reject $\theta_{\rm sky} = \theta_0$, and θ_0 can be taken as an upper limit to sky fluctuations.

The dotted curves in Figure 15 show how the limits on $\theta_{\rm sky}$ found by this method depend on the assumed errors when errors are overestimated ($\chi_\nu^2 < 1$). For each value of N the power is chosen to be the power of the likelihood ratio test in its "normal" form when $\chi_\nu^2 = 1$. For example, $\beta_{N=7} = 0.57$, $\beta_{N=15} = 0.55$, and $\lim_{N\to\infty} \beta = 0.5$. Such a modified likelihood ratio test is clearly safe for $\chi_\nu^2 > 1$, in the sense that it does not give misleadingly low limits.

The OVRO data set has roughly equal measurement errors for all fields and $\chi_\nu^2 \approx 1$, thus we obtain limits from the two statistical methods that are in close agreement: $\theta_{\rm sky} < 58$ μK (95% confidence) from a Bayesian analysis with a prior density uniform in $\theta_{\rm sky}$ for $\theta_{\rm sky}^2 \geq 0$, and $\theta_{\rm sky} < 52$ μK (95% confidence, $\beta = 0.72$) from a UMP likelihood ratio test. We will use 58 μK as our 95% limit.

b) Limits on the Temperature Autocorrelation Function

The statistical properties of a Gaussian random field, which we assume describes the microwave background fluctuations, are fully specified by the two-point correlation function. Over small angles, we can use the familiar Fourier expansion in rectangular coordinates for $T(x, y)$, the radiation temperature on the sky. The corresponding correlation function,

$$C(\phi) = \langle T(\boldsymbol{x}_1) T(\boldsymbol{x}_2) \rangle \,, \tag{6}$$

depends only on the angular distance $\phi \equiv |\boldsymbol{x}_1 - \boldsymbol{x}_2|$ between sampled points.

For physically reasonable fields, the maximum values of $C(\phi)$ occurs at $\phi = 0$, with zero first derivative. The *coherence angle* is defined by

$$\phi_c \equiv \left[-\frac{C(0)}{C''(0)} \right]^{1/2} .$$

Many examples of autocorrelation functions computed for models of interest can be found in the literature (e.g., Vittorio and Silk 1984; Vishniac 1987; Bond and Efstathiou 1987). Baryonic models with $0.1 \leq \Omega \leq 1$ and standard recombination (Peebles 1968) typically have coherence angles of 4′–8′ for adiabatic fluctutations (Gouda, private communication) and 2′–12′ for isocurvature modes (Efstathiou and Bond 1987). Models with lower values of Ω have smaller ϕ_c but larger amplitudes, as required by normalization with the observed galaxy-galaxy correlation function. Cold dark matter (CDM) models with $\Omega = 1$ have coherence angles of $\sim 10'$ for adiabatic fluctuation modes and 50′ for isocurvature modes (Bond and Efstathiou 1987). The effect of early reionization is to erase the fluctuations on small scales, effectively increasing ϕ_c and reducing the amplitude $C(0)^{1/2}$ (see, e.g., Efstathiou and Bond 1987), while generating new fluctuations on intermediate and small scales. Seond-order effects over the extended last scattering surface appear with amplitude $\leq 10^{-5}$ on scales $\phi_c \approx 1'.5$ (CDM) and 15′ (hot dark matter) for $\Omega = 1$ adiabatic models (Vishniac 1987). In isocurvature scenarios, for which early reionization or nonstandard recombination is most plausible, the reimposed perturbations appear with $C(0)^{1/2} \approx 10^{-5}$, again on arcminute scales (Efstathiou 1988).

In practice, we observe the sky with an antenna whose response pattern $B(x, y)$ (normalized to unit power over 4π sr) has nonzero width. The measured sky temperature as a function of position is the cross-correlation of the true temperature with the beam pattern,

$$T_{\rm obs}(x, y) = T \star B \,.$$

For a symmetric real beam the autocorrelation theorem implies

$$C_{\rm obs}(\phi) = C \star C_{\rm beam} \,, \tag{7}$$

where $C_{\rm beam}(\phi) \equiv B \star B$.

If, as is often the case, the antenna beams are well-represented by a circularly symmetric Gaussian with dispersion $\phi_0 = 0.4247 \phi_{\rm FWHM}$, equation (7) becomes

$$C(\phi_0, \phi) = C(\phi) \star \frac{1}{4\pi\phi_0^2} \exp\left(-\frac{\phi^2}{4\phi_0^2} \right), \tag{8}$$

where we write $C(\phi_0, \phi)$ in place of $C_{\rm obs}$ as an explicit reminder of the Gaussian approximation. This expansion is analytically tractable for many $C(\phi)$. For the OVRO 40 m telescope, $\phi_{\rm FWHM} = 108''$, so $\phi_0 = 0'.77$.

For switching experiments, one must cross-correlate the smeared temperature field $T_{\rm obs}(x, y)$ with the sampling function $S(x, y)$. For an idealized double switching scheme with three Gaussian beams separated by ϕ_S (see Fig. 3 and eq. [3]), the expected sky variance is

$$\begin{aligned} \langle \Delta T^2 \rangle &= \langle [T_M - \tfrac{1}{2}(T_{R1} + T_{R2})]^2 \rangle \\ &= \tfrac{3}{2} \langle T_{\rm obs}(\boldsymbol{x})^2 \rangle - 2 \langle T_{\rm obs}(\boldsymbol{x}_M) T_{\rm obs}(\boldsymbol{x}_R) \rangle \\ &\quad + \tfrac{1}{2} \langle T_{\rm obs}(\boldsymbol{x}_{R1}) T_{\rm obs}(\boldsymbol{x}_{R2}) \rangle \,, \end{aligned} \tag{9}$$

where $|\boldsymbol{x}_M - \boldsymbol{x}_R| = \phi_S$, and $|\boldsymbol{x}_{R1} - \boldsymbol{x}_{R2}| = 2\phi_S$. Combining equations (6), (8), and (9), we obtain

$$\langle \Delta T^2 \rangle = \tfrac{3}{2} C(\phi_0, 0) - 2C(\phi_0, \phi_S) + \tfrac{1}{2} C(\phi_0, 2\phi_S) \,. \tag{10}$$

On the 40 m telescope, $\phi_S = 7'.15$. Analytic tractability is often lost when the details of real experiments must be included; however, as discussed in § VIII*c* below, equation (10) turns out to be quite accurate for the OVRO measurements. (See also Boynton 1980 for discussion of the manipulation of autocorrelation functions.)

If $\phi_c \ll \phi_S$, equation (10) reduces to

$$\langle \Delta T^2 \rangle \approx \tfrac{3}{2} C(\phi_0, 0) \,.$$

If in addition $\phi_c \gg \phi_0$, then beam smearing is unimportant, and

$$\langle \Delta T^2 \rangle \approx \tfrac{3}{2} C(0) . \quad (11)$$

In real experiments, the ratio ϕ_S/ϕ_0 must usually be restricted in order to minimize certain systematic errors, and both inequalities cannot be satisfied simultaneously. Nevertheless, equation (11) is often a good approximation for $\phi_c \approx \sqrt{\phi_0 \phi_S}$, and it has been used widely in the past to derive limits from switching experiments (e.g., Uson and Wilkinson 1984*a*, *b*; Boynton and Partridge 1973). In § VIII*a* we derived an upper limit on the sky dispersion from the OVRO measurements, assuming Gaussian fluctuations, of $\theta_{\rm sky} < 58$ μK. In the approximation of equation (11), then,

$$\frac{\text{“}\delta T\text{”}}{T} = \frac{C^{1/2}(0)}{T} = \frac{\sqrt{2/3}\,\theta_{\rm sky}}{2.78} < 1.7 \times 10^{-5} .$$

As noted above, many models have large-scale power, with coherence angles outside the range where equation (11) is a reasonable approximation. To put our results in a form that can be compared with a wide range of model predictions, we assume that the temperature autocorrelation function can be approximated by a Gaussian,

$$C(\phi) = C_0 \exp\left(-\frac{\phi^2}{2\phi_c^2}\right).$$

In this case equation (8) becomes

$$C(\phi_0, \phi) = C_0 \frac{\phi_c^2}{2\phi_0^2 + \phi_c^2} \exp\left[-\frac{\phi^2}{2(2\phi_0^2 + \phi_c^2)}\right]. \quad (12)$$

This approximation has been used only occasionally in the past to present results of observations (Davies *et al.* 1987), but usually produces results in good agreement with those obtained from the detailed model autocorrelation functions. Figure 16 shows the OVRO limits on C_0 for this model as a function of ϕ_c, obtained by combining equations (10) and (12), and using the OVRO experimental limit of 58 μK derived in the previous section. Note the reduced sensitivity to fluctuations with coherence angles smaller than the beam size due to averaging over many independent fluctuations, and to fluctuations with large-scale power where we sample only the curvature of a single fluctuation.

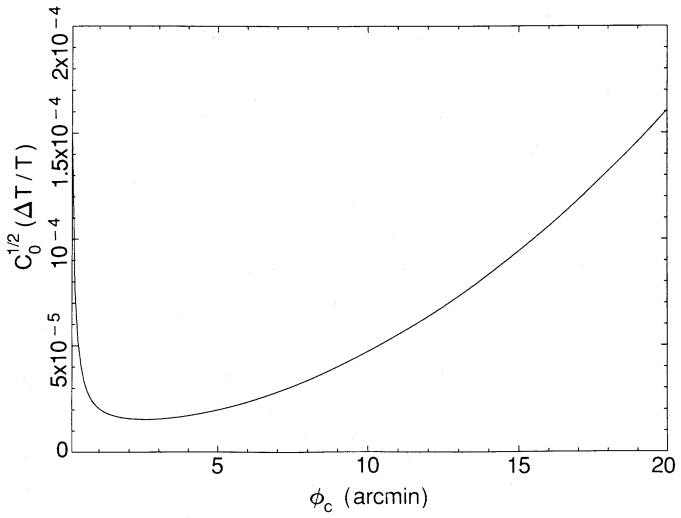

FIG. 16.—Limits on $\Delta T/T$ for a Gaussian correlation function, assuming Gaussian beams in both center and reference fields, for $\theta_{\rm sky} = 58$ μK.

In the case of non-Gaussian statistics for the temperature field, such as those produced by models involving cosmic strings (Ostriker, Thompson, and Witten 1986; Bouchet, Bennett, and Stebbins 1988) and decaying massive particles (Daly 1987), the correlation function and hence $\theta_{\rm sky}$ is a poor discriminator between theories. Such models predict relatively rare signals of large amplitude. As discussed in the Appendix, the proper method to search for these fluctuations is to cover a large portion of sky less deeply, trading sensitivity for area. The effective beam area of the 40 m at 20 GHz is $2\pi\phi_0^2 = 3.72$ arcmin^2, giving a total of 61 arcmin^2 for the eight fields and 16 reference positions. More sophisticated analyses such as $P(D)$ distribution tests used in source count studies (Scheuer 1957, Condon 1974) can also provide useful limits on non-Gaussian source distributions, especially those of power-law form. Note that by increasing the search area the survey becomes more susceptible to contamination by discrete radio sources, which have just the power-law–number flux distribution to which this test is particularly sensitive. In this case, it is necessary to interlock the main and reference fields or to map a contiguous region of the sky to allow discrimination and identification of features.

c) Model Comparisons and Corrections for the True Beam

If a model of MWB fluctuations has Gaussian fluctuations, the results can be given in terms of an autocorrelation function. Then equation (7) shows how to compare the model with our observations. The beam correction depends on both the autocorrelation function and the detailed shape of the beam shown in Figure 4. Clearly, an analytic approximation for Figure 4 would be useful for this purpose. Since the instantaneous beams of the 40 m are reasonably close to Gaussian, the simplest function that could represent Figure 4 consists of one positive Gaussian of unit height straddled by two negative Gaussians half as high and 7.′15 away, all of FWHM 1.′8.

We have estimated the error that would be introduced by using this simplified beam in comparisons with models, in the following way. Bond and Efstathiou (1987) calculated the rms angular power spectrum for two CDM models. Using these power spectra (supplied by Bond), we made several realizations of maps of the microwave background radiation according to each model. We convolved these maps with two “beams”: the effective beam of Figure 4, based on measurements of the true telescope beams; and the Gaussian approximation mentioned above. The value of the convolution at a given point is just the mean temperature ΔT that we would measure on the model sky in the absence of instrumental noise, so that the rms value of the convolved map is just $\theta_{\rm sky}$. For both models, the differences between the results with the true beam and the Gaussian approximation were less than 2%.

This close agreement depends to some extent on the models, and we cannot generalize the results to arbitrary power spectra. However, any model whose power spectrum is not drastically different from those of the Bond and Efstathiou models can be compared with the OVRO results using a simple three-Guassian analytic beam without fear of significant error.

We can test the independence of our observed fields in a similar fashion, by comparing the convolution with a single effective beam to the convolution with seven beams appropriately spaced around a 1° ring. For both adiabatic and iso-

curvature models the convolutions differed by the factor of $7^{1/2}$ expected if the measurements are truly independent.

IX. COMPARISON WITH OTHER OBSERVATIONS

A number of sensitive observations on the isotropy of the microwave background radiation have been made. It is often difficult to compare results made with different instruments, on different angular scales; however, if the autocorrelation function of the background fluctuations is smooth, anisotropy limits from measurements on one angular scale can be extrapolated to other angular scales (e.g., Fig. 16). In this section we discuss the most recent results on four different angular scales.

On scales of 1′ or less, the most sensitive published measurements are the VLA results of Martin and Partridge (1988) and Fomalont *et al.* (1988). Martin and Partridge report a detection of fluctuations at levels of 1.7×10^{-4} and 1.3×10^{-4} on scales of 36″–160″ and 18″–80″. Fomalont *et al.* derive a 95% confidence limit of $\delta T/T < 6 \times 10^{-5}$ on a scale of 1′, and other limits on smaller scales as shown in Figure 17. The results and techniques of the two groups are compared in some detail by Partridge (1989), who concludes that the cause of this apparent contradiction is not fully understood. As will be seen below, under certain assumptions the OVRO results favor interpretation of the VLA measurements as upper limits rather than detections, but our comparisons should not be taken as a resolution of the interesting questions raised by Partridge (1989).

On scales of a few arc minutes, the most sensitive previous measurement is that of Uson and Wilkinson (1984*a*, *b*, *c*), who used an experimental setup similar to our own. They found that $\delta T/T < 2.1 \times 10^{-5}$ at the 95% confidence level, using a likelihood ratio test as described in § VIII*a*(ii). Unfortunately, the power of the likelihood ratio test applied to their measurements is only 0.13, a result of the fact that their measurements were closer to zero than expected from the measurement errors. The Bayesian analysis of § VIII*a*(i) applied to their data (given in Lasenby 1988) yields $\delta T/T < 3.8 \times 10^{-5}$ for a uniform prior distribution. However, the simulations described in § VIII*a*(iii) suggest that this is still an underestimate, given that $\chi_\nu^2 = 0.67$ for their data, and that 4.7×10^{-5} is more consistent with their error estimates.

On scales from a few arc minutes up to 2°.5, the lowest published upper limits are those of Parijskij and his coworkers (Parijskij 1973*a*, *b*; Parijskij, Petrov, and Cherkov 1977; Berlin *et al.* 1983, 1984). Their early results range from 1.3×10^{-5} to 8.0×10^{-5}. These have been converted to 95% confidence upper limits by Partridge (1980*a*, *b*; 1983), but not corrected for possible errors in statistical analysis. Including such factors, Lasenby (1981) calculates 95% confidence upper limits ranging from 5.4×10^{-5} on a scale of 75′ to 1×10^{-4} on a scale of 10′. Recent preliminary results from this group (Berlin *et al.* 1983) give $\delta T/T < 1 \times 10^{-5}$ (1 σ level) on scales from 4′.5 to 9′.5, and $\delta T/T < 3 \times 10^{-5}$ (1 σ level) on a scale of 1°. These observations were made with the Ratan 600 m telescope at a frequency of 3.9 GHz. Observations at this frequency with this telescope require substantial corrections for background sources. Amirkhanyan (1987) has estimated that due to the effects of confusion the reported upper limits are optimistic by a factor of 10, but his analysis has been challenged by Parijskij, Petrov, and Cherkov (1987). As yet we do not have enough details of the observing and analysis procedures to compare these results directly with our own, and we do not, therefore, consider them further here.

On a scale of 8°.3, with a beam dispersion of $\phi_0 = 3°.5$, Davies *et al.* (1987) report the detection of anisotropies at a level of 3.7×10^{-5}, corresponding to a triple-beam $\theta_{\rm sky}$ of 2.9×10^{-5}. They are now following up this 10.4 GHz detection with observations at other frequencies to determine whether this anisotropy is intrinsic or galactic in origin. The Relikt experiment on the *Prognoz 9* satellite yielded a residual rms temperature fluctuation of 0.2 mK after convolution with a 7° (FWHM) Gaussian (Klypin *et al.* 1987). Analysis of these data based on an assumed fractal perturbation spectrum gives $\delta T/T < 5.6 \times 10^{-6}$ on a scale of 6°. However, the Relikt observations are consistent with the Davies *et al.* result for a power-law spectrum with index $n > 1$ (Scaramella and Vittorio 1988).

In Figure 17 we compare our results with those of Fomalont *et al.* and Davies *et al.*, assuming a Gaussian autocorrelation function. The curve for the Davies *et al.* observations was computed in the same manner as that in Figure 16 for the OVRO limit, also reproduced in Figure 17. The appropriate expression to allow inclusion of the VLA results of Fomalont *et al.* was obtained using the autocorrelation function $C(\theta_0, \theta)$ (see § VIII*b*) for a synthesized beam of dispersion θ_0, the Fourier transform relationship between the autocorrelation function and the power spectrum, and the effective truncation of the measured power spectrum for interferometer spacings less than the VLA telescope diameter. Curves are plotted for $\theta_{\rm sky} < 8.3 \times 10^{-4}$, 1.2×10^{-4}, 7.8×10^{-5}, and 5.8×10^{-5}, with synthesized beams of FWHM 12″, 18″, 30″, and 60″, respectively. The OVRO results provide the most stringent limits on $C_0^{1/2}$ on all scales smaller than 26′, while those of Davies *et al.* provide the most stringent limits on larger scales. From Figure 17, it is clear that if the Davies *et al.* anisotropy turns out to be intrinsic to the microwave background, and the autocorrelation function is Gaussian, then $\phi_c > 26'$.

X. DISCUSSION

Proposed theories of galaxy formation can be classified as linear perturbation theories, which rely on linear growth of small density fluctuations until the density contrast approaches unity and objects condense out of the Hubble flow (Lifschitz 1946), or nonlinear theories in which galaxy formation is driven by something other than the gradual growth of perturbations. Most linear theories assume either adiabatic fluctuations (e.g., Peebles and Yu 1970; Sunyaev and Zel'dovich 1972; Doroshkevich, Zel'dovich and Sunyaev 1978; Silk and Wilson 1980; Wilson and Silk 1981; Kodama and Sasaki 1986; Bond and Efstathiou 1987), or entropy fluctuations (e.g., Peebles 1974; Gott and Rees 1975; Silk and Wilson 1980; Vittorio and Silk 1984; Efstathiou and Bond 1986; Bond and Efstathiou 1987; Efstathiou and Bond 1987; Gouda, Sasaki, and Suto 1987). A variety of nonlinear mechanisms have been considered, including explosions of primordial stars (Doroshkevich, Zel'dovich, and Novikov 1967; Ikeuchi 1981; Ostriker and Cowie 1981), cosmic strings (Ostriker, Thompson, and Witten 1986; Stebbins 1988), and mock gravity (Hogan and White 1986). These scenarios are further complicated by the possibilities of early reionization (Hogan 1980, 1984; Kaiser 1984*a*; Ostriker and Vishniac 1986; Peebles 1987*a*; Vishniac 1987; Silk and Vittorio 1987; Efstathiou 1988), baryonic and nonbaryonic dark matter (White and Rees 1978; Vittorio and Silk 1984; Efstathiou and Bond 1986; Bond and Efstathiou 1987), and biased galaxy formation (Kaiser 1984*a*, *b*, 1986).

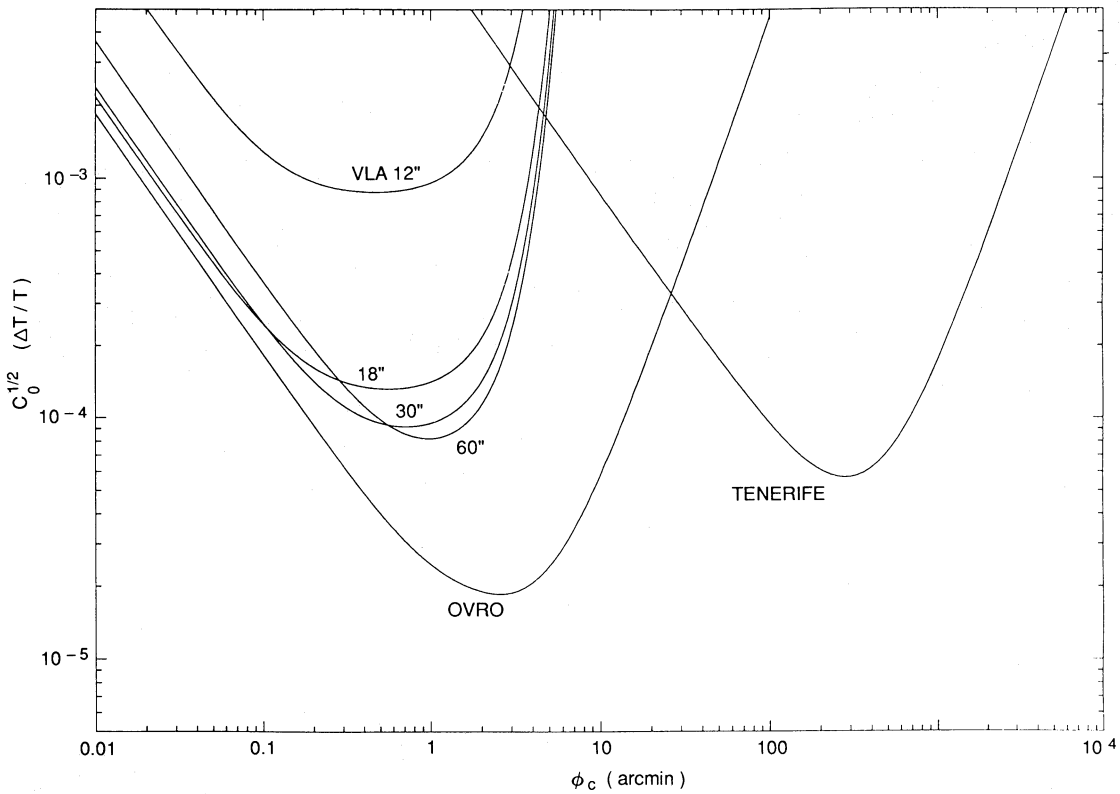

FIG. 17.—Comparison of limits placed on a Gaussian autocorrelation function by the OVRO measurements, with those placed by the measurements of Fomalont *et al.* (1988) and Davies *et al.* (1987).

For convenience, we will refer to models in which the dominant matter constituent of the universe is baryonic and the reionization of the intergalactic medium occurred comparatively recently (i.e., after the epoch at which it would obliterate intrinsic anisotropy on small scales) as "conventional" models. Models based on nonbaryonic matter, early reionization, biased galaxy formation or explosions will be termed "unconventional." No great significance is attached to these labels. It will be seen that the new anisotropy limits place interesting constraints on the conventional models, but do not strongly constrain unconventional models.

It is well known that the intergalactic medium is highly ionized back to the redshifts of distant quasars (see, e.g., Steidel and Sargent 1987 for a recent discussion). Intrinsic fluctuations are not erased by plasma at a given redshift on angular scales greater than the horizon scale at that redshift. It is easy to show that for all plausible values of the density parameter intrinsic fluctuations on angular scales of arc minutes will be erased only if reionization occurs before redshift 10. Since there is no evidence of the existence of compact objects at redshifts greater than 10, reionization earlier than that seems unlikely. Thus conventional scenarios embrace those models in which reionization of the intergalactic medium is due, for example, to ultraviolet radiation from quasars and occurs after redshift 5 (e.g., Donahue and Shull 1987).

The growing conflict between observations and theoretical predictions of conventional models has spurred interest in unconventional models. Numerous models, both linear and nonlinear, have invoked various combinations of early reionization, nonbaryonic matter and biased galaxy formation to reconcile this conflict. Many of these models have not been developed to the point where they make detailed predictions about expected levels of anisotropy, so they cannot yet be tested against our new limit. Those cases for which estimates have been given are discussed below.

There are many predictions of the temperature anisotropy of the microwave background radiation in the literature. As discussed in § VIII*b*, a precise comparison of experimental results with model predictions must take account of both beamwidth and sampling effects. This is possible only when the model autocorrelation function (or, equivalently, the power spectrum) is given. Unfortunately, models are often given in the literature without this information. Sometimes enough information is given to permit an approximate reconstruction of the autocorrelation function or power spectrum, and a fairly accurate prediction for our experimental arrangement can be determined.

Table 6 gives the levels of $\delta T/T \times 10^5$ predicted by various models after adjustment for the parameters of our observations. The models assume a power-law initial spectrum of the form $|\delta_k|^2 \propto k^n$, where δ represents the fluctuation of density or entropy as appropriate. The corrections that we have made are given in the footnotes. Values in the table can be compared directly with the OVRO limits of 2.1×10^{-5} (95%) or 4.6×10^{-5} (equivalent 3 σ).

Clearly, one should be cautious in accepting or rejecting models based on fine distinctions. Nevertheless, there are some clear results:

1. Most of the predictions of conventional adiabatic models are well above our equivalent 3 σ limit, and are therefore definitively ruled out. The lowest predictions of $\delta T/T$ are

TABLE 6

PUBLISHED MODEL PREDICTIONS OF "$\delta T/T$" $\times 10^5$, CORRECTED WHERE POSSIBLE FOR OVRO BEAM AND BEAM SEPARATION†

MODEL PARAMETERS‡						FLUCTUATIONS†					
Ω	Ω_b	Ω_{nb}	$\Omega_{vac}h$		n	Adiabatic	Adiabatic + Early Reionization	Isocurvature	Isocurvature + Early Reionization	Explosions	COMMENTS
.1	.1	0	0	.5	0				1.5 *c* A		
.1	.1	0	0	.5	−1	72 *k* G			3.2 *c* A		
.1	.1	0	0	.5	0	100 *k* G			0.1 *d* D		
.1	.1	0	0	.5	−1				0.9 *d* D		
.1	.1	0	0	.5	+1	87 *k* G		55 *e* E			
.1	.1	0	0	1	+1			80 *e* E			
.1	.1	0	0	.5	?					>5.2 *f* F	
.1	.1	0	0	1	?					> 10 *f* F	
.2	.2	0	0	.5	0			10.0 *c* I	0.5 *c* I		
.2	.2	0	0	.5	−1			16.5 *c* I	1.7 *c* I		
.2	.2	0	0	.5	0				2.3 *c* A		$z < 200$
.2	.2	0	0	.5	−1				2.3 *c* A		$z < 200$
.2	.2	0	0	.5	0				1.2 *c* A		$z < 200$, $\chi_e = 0.1$
.2	.2	0	0	.5	−1				2.7 *c* A		$z < 200$, $\chi_e = 0.1$
.2	.2	0	0	1	0			10 *c* A			
.2	.2	0	0	1	−1			17 *c* A			
.4	.4	0	0	.5	0				0.9 *d* D		
.4	.4	0	0	.5	−1				1.7 *d* D		
.4	.4	0	0	.5	+1			14 *e* E			
1	1	0	0	.5	0				3.0 *d* D		
1	1	0	0	.5	0	4.3 *l* G		11.0 *c* I	8.5 *c* A		
1	1	0	0	.5	−1	2.4 *l* G		11.1 *c* I	3.0 *c* A		
1	1	0	0	.5	−1				1.4 *d* D		
1	1	0	0	.5	−2			5.3 *c* A			
1	1	0	0	.5	+1	4.4 *l* G		1.9 *e* E			
1	1	0	0	1	+1			4.0 *e* E			
.2	.03	.17	0	.5	+1	13 *a* I					
.2	.1	.1	0	.75	+1	17 *a* B					
.2	0	.2	0	.5	+1	17 *i* G					
.4	0	.4	0	.5	+1	5.4 *i* G					
.4	0	.4	0	1	+1	2.0 *i* G					
1	0	1	0	.5	+1	1.3 *i* G					
1	0	1	0	1	+1	0.7 *i* G					
.4	.03	.37	0	.5	0	4.4 *j* E					Biased
.4	.03	.37	0	.5	+1	3.8 *j* E					Biased
1	.03	.97	0	.5	0	0.9 *j* E					Biased
1	.03	.97	0	.5	+1	0.9 *j* E					Biased
1	.03	.97	0	.5	+1	0.5 *b* B					Biased 1.7
1	.1	.9	0	.5	+1	0.7 *a* B	.03 *m* I				Biased 1.7
1	.2	.8	0	.4	+1	0.9 *a* I					Biased 1.7
1	.5	.5	0	.5	+1	1.7 *a* I					Biased 1.7
1	.1	.9	0	.5	...					> 15 *f* F	
1	.1	.9	0	1	...					> 30 *f* F	
1	.1	.9	0	.5	+1		0.5 *h* G				
1	.1	.9	0	1	+1		0.8 *h* G				
1	.1	.9	0	.5	+1		0.1 *h* H				Biased
1	.1	.9	0	1	+1		0.1 *h* H				Biased
1	.1	.9	0	.5	0		1.3 *h* G				Biased HDM
1	.1	.9	0	1	0		1.6 *h* G				Biased HDM
1	.1	.9	0	.5	+1	2.5 *a* I					Antibiased massive neutrinos
1	.03	.17	.8	.75	+1	1.7 *m* I					$\Lambda \neq 0$
1	.03	.17	.8	.5	+1	3.5 *a* I					$\Lambda \neq 0$
1	.1	.1	.8	.5	+1	1.5 *a* B					$\Lambda \neq 0$
1	~0	~1	0	.75	?			0.4 *m* I			Biased 1.7 Axions
1	~0	~1	0	.5	+1	3.6 *g* G					Massive neutrinos

obtained for high-density models ($\Omega_b \approx 1$). However, even these are well above our 95% confidence limit.

2. Conventional isocurvature models by Efstathiou and Bond (1987) and Gouda, Sasaki, and Suto (1987) also produce fluctuations above our 95% limit for $\Omega_b \leq 1$. Models with $\Omega_b \leq 0.8$ are difficult to reconcile with our present limit, and an improvement of a factor of 2 in our limit would provide strong constraints over a large region of the Ω_b–n plane.

3. Nonbaryonic models with early reionization predict anisotropy levels up to a factor of 3 below our present limit.

4. Many isocurvature baryonic models with early reionization predict anisotropy levels slightly below our limit, but much of the Ω_b–n plane would be excluded if no anisotropy were detected at half the present limit (Efstathiou 1988).

5. The lowest predictions come from models with biased galaxy formation, nonbaryonic matter and early reionization, and are as much as a factor of 10 below our present sensitivity limit.

6. Some massive neutrino models are excluded (Silk 1984), but not all (Bond 1988).

Based on most theories suggested thus far, a modest improvement in the present sensitivity level would lead to the detection of anisotropy. We regard it as encouraging that most theories of galaxy formation, including not only linear theories with early reionization or nonbaryonic matter, but also nonlinear theories, such as those based on superconducting cosmic strings (Ostriker and Thompson 1987), are within reach with only modest extensions of present techniques. If no anisotropy is detected within a factor of 3 of the present limits, most present theories of galaxy formation will be in jeopardy. Possibly those invoking nonstandard reionization and nonbaryonic matter and biased galaxy formation will still be tenable, particularly if further relaxation of the normalization requirements is justified.

We believe that sensitivity 3 or 4 times better than that of the present work should be achievable on arcminute scales with ground based observations at centimeter wavelengths. Discrete source confusion and low-level systematic errors are likely to be the limiting factors. If no anisotropy is detected at this level, alternative methods will probably be needed. Two promising possibilities are instruments designed to image the microwave background radiation on angular scales up to 20°, and a space antenna designed to measure microwave background radiation anisotropy on angular scales down to $\frac{1}{2}°$ with a sensitivity of 1 μK.

We would like to thank R. Bond and N. Kaiser, organizers of a workshop at the Canadian Institute of Theoretical Astrophysics, and the other participants in that workshop, for many useful discussions about the microwave background radiation. E. Turner, S. Boughn, D. Cottingham, A. Lasenby, and N. Kaiser contributed significantly to our statistical education. We are grateful to R. Bond and G. Efstathiou for sharing the results of their calculations before publication, and to P. Crane for the VLA observations of NCP 9 on short notice. We thank C. Hogan and B. Partridge for useful criticisms of an early version of this paper, and G. Bernstein, R. Bond, P. Boynton, G. Efstathiou, A. Lasenby, J. Uson, and D. Wilkinson for useful discussions. This work would not have been possible without the help of a number of our colleagures in developing the maser, particularly R. Moore, and R. Clauss, G. Resch, S. Petty, and D. Neff at the Jet Propulsion Laboratory. We are particularly indebted to R. Vogt for this help and encouragement at a crucial time, without which the Microwave Background Program at the Owens Valley Radio Observatory would likely have foundered, and to M. Cohen for his support of this program over many years. We also thank the staff of the OVRO, particukarly T. Seling, M. Hodges, C. Giovanine, W.

NOTES TO TABLE 6

† The letters following the predictions refer to the lists of references and corrections given in the notes. Values in the table should be compared directly with the OVRO limits of 2.1×10^{-5} (95%) or 4.6×10^{-5} (equivalent 3 σ).

‡ Total mass-energy density; its constituents in baryons; nonbaryons; and the vacuum; the Hubble constant in units of 100 km s^{-1} Mpc^{-1}; and the spectral index of the initial fluctuations, given by $|\delta_k^2| \propto k^{3+n}$.

REFERENCES.—Key to references following the predictions:

a Bardeen, Bond, and Efstathiou 1987.
b Bond 1988.
c Efstathiou and Bond 1987.
d Efstathiou 1988.
e Gouda, Sasaki, and Suto 1987.
f Hogan 1984.
g Silk 1986.
h Vishniac 1987.
i Vittorio and Silk 1984.
j Vittorio, Materrese, and Lucchin 1988.
k Wilson 1983.
l Wilson and Silk 1981.
m Bond 1988 (private communication).

CORRECTIONS.—Model predictions are often given for particular beam sizes, beam separations, and switching schemes. Where possible we have multiplied published values by a factor f so that the predictions correspond to the OVRO beam arrangement and switching scheme.

A Prediction given for Uson and Wilkinson beam size and separation. Corrected to OVRO using ϕ_c given by or estimated from reference. $1.1 \leq f \leq 2.5$.
B Predictions from reference *a* were modified for OVRO observations in reference *b*. Predictions from reference *b* given for OVRO observations. No correction needed.
C Prediction given for Uson and Wilkinson experiment. Corrected by $f = 2$ under the assumption $\phi_c \geq 6'$ (note $\phi_c \approx 12'$ for the $\Omega = 1, \Omega_b = 0.03, h = 0.75$ CDM models of Bond and Efstathiou 1987). If $\phi_c \geq 20'$, then $f = 2.5$.
D Prediction given for OVRO. No correction needed.
E Prediction given for OVRO ϕ_s, but $\phi_0 = 0.64$ (the Uson and Wilkinson value). No correction made.
F Fluctuations have no angular scale. No correction made.
G Prediction given in terms of ϕ_s, but $\phi_0 = 1'.5$. Not corrected to $\phi_0 = 0'.78$, the OVRO value. Corrected from single to double switching if necessary.
H Prediction given in terms of a biasing factor. Biasing at 2.7 σ peaks assumed, $f = 0.17$. Otherwise same as G.
I Prediction for OVRO from reference *m*. No correction needed.

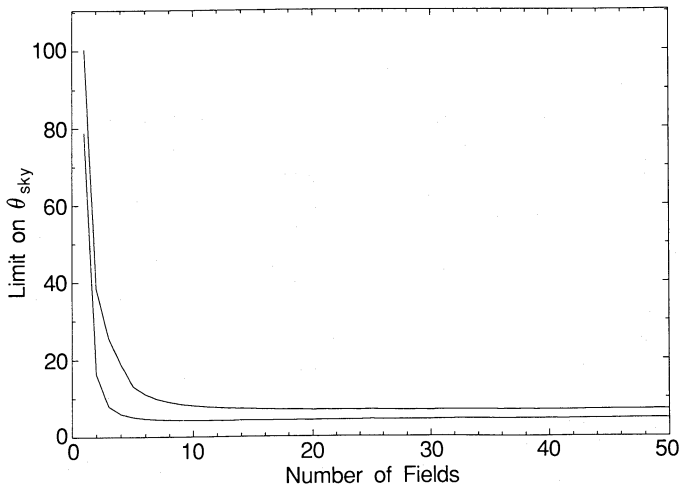

FIG. 18.—Relative upper limits that would be placed on $\delta T/T$ for measurements with a given total integration time divided among N fields, assuming Gaussian fluctuations and using the Bayesian method of § VIII*a*(i) with a uniform prior distribution. The upper curve gives 0.9987 probability, the lower 0.95.

Hutton, W. Schaal, L. Wiley, and C. Lackore. We gratefully acknowledge the support of the National Science Foundation, the James Irvine Foundation, the IBM Research Fund, and the Caltech President's Fund. This work has been supported by the National Science Foundation under grants AST 79-13249, AST 80-24119, AST 82-10259, AST 85-09822, and AST 86-10693.

This paper is dedicated to the memory of Alan Moffet, who contributed much to this work before his untimely death on 1987 August 20.

APPENDIX

In this appendix, we discuss the optimum number of fields to observe, given a limited available integration time τ. The optimum number depends on the relationship between integration time and sensitivity, the spectrum of fluctuations we are trying to measure, and on various instrumental factors.

For the reasons given in § V, we are confident that the sensitivity of our measurements with the 40 m telescope is approximately proportional to the square root of integration time. Thus if a total integration time τ is split between N fields, and σ is the error for $N = 1$, the error σ_N in time τ/N is given by $\sigma_N^2 = N\sigma^2$.

The fluctuation spectrum, on the other hand, is unknown, so we can only answer the question conditionally. For Gaussian fluctuations, the answer is best found using the Bayesian analysis of § VIII*a*(i). As before, we assume that the fluctuations have dispersion θ_{sky}. With a uniform prior distribution [$p(\theta_{sky}) = \text{constant}$], $p(\theta_{sky} \mid \{\Delta T_i\}) \propto L(\{\Delta T_i\} \mid \theta_{sky})$. To estimate the limit that an N-field experiment could place on θ_{sky}, we assume that the measured values ΔT_i are determined entirely by the measurement errors, that is, $\theta_{sky} = 0$. Then the ΔT_i are normally distributed with variance $N\sigma^2$. From the posterior distribution, we determine θ^*_{sky} such that $\int_0^{\theta^*_{sky}} p(\theta_{sky} \mid \{\Delta T_i\}) d\theta_{sky} = c$, where $c = 0.95$, 0.9987, or some other favorite value. θ^*_{sky} is the 95% or equivalent 3 σ upper limit that we would place on θ_{sky} from our measurements. Figure 18 shows that as N increases, θ^*_{sky} decreases sharply at first, then levels out with an extremely broad minimum at $N \approx 14$ for $c = 0.95$ and $N \approx 25$ for $c = 0.9987$. For $c = 0.95$, θ^*_{sky} is almost constant for $N \gtrsim 10$, while for $c = 0.9987$, θ^*_{sky} is almost constant for $N \gtrsim 17$.

Thus, for Gaussian fluctuations and thermal noise, there is no clear choice of N, except that it not be too small. However, three additional considerations favor moderate values of N over large ones. First, the smaller N, the closer to the celestial pole the fields can lie, minimizing systematic errors from differential ground pickup. Second, for moderate values of N the errors in each field are small. Low-level systematic effects that are detectable in measurements of individual fields with small errors might be undetectable in observations of many fields with much larger individual errors, yet still distort the overall result. Finally, when N is large fewer observations will be made of each individual field, making it difficult to estimate the errors, particularly those due to long-term atmospheric fluctuations. As we have shown in § VIII*a*(iii), both under- and overestimates of the true errors have serious consequences in the statistical analysis of data.

Similar calculations could be done for any assumed distribution of sky fluctuations. Few methods with non-Gaussian fluctuations have been proposed, so we restrict ourselves here to a general consideration. Suppose that the density of fluctuations of temperature ΔT is given by $\rho(\Delta T) \propto \Delta T^\beta$. If $\beta = -2$, the probability of finding *one* source in N fields at some fixed multiple of the noise level σ (proportional to $N^{1/2}$ from above) is independent of N. For $\beta > -2$, large fluctuations will be detected with higher probability in an experiment with large N.

REFERENCES

Aaronson, M., Bothun, G., Mould, J., Huchra, J., Schommer, R. A., and Cornell, M. E. 1986, *Ap. J.*, **302**, 536.
Aaronson, M., Huchra, J., Mould, J., Schechter, P. L., and Tully, B. 1982, *Ap. J.*, **258**, 64.
Amirkhanyan, V. R. 1987, *Soob. Spec. Astrofiz. Obs.*, **53**, 96.
Baars, J. W. M., Genzel, R., Pauliny-Toth, I. I. K., and Witzel, A. 1977, *Astr. Ap.*, **61**, 99.
Bahcall, N. A. 1987, *Comments Ap.*, **11**, 283.

Bahcall, N. A., and Soneira, R. M. 1982, *Ap. J.*, **262**, 419.
Bardeen, J. M., Bond, J. R., and Efstathiou, G. 1987, *Ap. J.*, **321**, 28.
Bardeen, J. M., Steinhardt, P., and Turner, M. S. 1983, *Phys. Rev. D*, **28**, 679.
Berger, J. O. 1985, *Statistical Decision Theory and Bayesian Analysis* (New York: Springer).
Berger, J. O., and Wolpert, R. L. 1984, *The Likelihood Principle* (Hayward: Institute of Mathematical Statistics).
Berlin, A. B., Bulaenko, E. V., Vitkovsky, V. V., Kononov, V. K., Parijskij, Yu. N., and Petrov, Z. E. 1983, in *IAU Symposium 104, The Early Evolution of the Universe and Its Present Structure*, ed. G. Abell and G. Chincarini (Dordrecht: Reidel), p. 121.
Berlin, A. B., Gassanov, L. G., Gol'nev, V. Ya., Korol'kov, D. V., and Parijskij, Yu. N. 1984, *Soob. Spec. Astrofiz. Obs.*, **41**, 5.
Bertschinger, E., and Watts, P. W. 1988, *Ap. J.*, **328**, 23.
Birkinshaw, M., Gull, S. F., and Hardebeck, H. E. 1984, *Nature*, **309**, 34.
Bond, J. R. 1988, in *IAU Symposium 130, Evolution of Large Scale Structures in the Universe*, ed. J. Audouze, M. C. Pelletan, and A. Szalay (Dordrecht: Kluwer), p. 93.
Bond, J. R., and Efstathiou, G. 1987, *M.N.R.A.S.*, **226**, 655.
Bond, J. R., Szalay, A. S., and Silk, J. 1988, *Ap. J.*, **324**, 627.
Bouchet, F. R., Bennett, D. P., and Stebbins, A. 1988, *Nature*, **335**, 410.
Boynton, P. E. 1980, in *IAU Symposium 92, Objects at High Redshift*, ed. G. O. Abell and P. J. E. Peebles (Dordrecht: Reidel), p. 293.
Boynton, P. E., and Partridge, R. B. 1973, *Ap. J.*, **181**, 243.
Brandenberger, R., Albrecht, A., and Turok, N. 1986, *Nucl. Phys. B*, **277**, 605.
Condon, J. J. 1974, *Ap. J.*, **188**, 279.
Daly, R. A. 1987, *Ap. J.*, **322**, 20.
Danese, L., De Zotti, G., and Mandolesi, N. 1983, *Astr. Ap.*, **121**, 114.
Davies, R. D., Lasenby, A. N., Watson, R. A., Daintree, E. J., Hopkins, J., Beckman, J., Sanchez-Almeida, J., and Rebolo, R. 1987, *Nature*, **326**, 462.
Donahue, M., and Shull, M. J. 1987, *Ap. J. (Letters)*, **323**, L1.
Donnelly, R. H., Partridge, R. B., and Windhorst, R. A. 1987, *Ap. J.*, **321**, 94.
Doroshkevich, A. G., Zel'dovich, Ya. B., and Novikov, I. D. 1967, *Soviet Astr.*, **11**, 233.
Doroshkevich, A. G., Zel'dovich, Ya. B., and Sunyaev, R. A. 1978, *Soviet Astr.*, **22**, 523.
Dressler, A., Faber, S. M., Burstein, D., Davies, R. L., Lynden-Bell, D., Terlevich, R. J., and Wegner, G. 1987, *Ap. J. (Letters)*, **313**, L37.
Edwards, A. W. F. 1984, *Likelihood* (Cambridge: Cambridge University Press).
Efstathiou, G. 1988, in *Large-Scale Motions in the Universe*, ed. V. C. Rubin and G. V. Coyne (Princeton: Princeton University Press), p. 299.
Efstathiou, G., and Bond, J. R. 1986, *M.N.R.A.S.*, **218**, 103.
———. 1987, *M.N.R.A.S.*, **227**, 33P.
Fomalont, E. B., Kellermann, K. I., Anderson, M. C., Weistrop, D., Wall, J. V., Windhorst, R. A., and Kristian, J. A. 1988, *A.J.*, **96**, 1187.
Fomalont, E. B., Kellermann, K. I., and Wall, J. V. 1984, *Ap. J. (Letters)*, **277**, L23.
Franceschini, A., Toffolatti, L., Danese, L., and De Zotti, G. 1989, *Ap. J.*, **344**, 35.
Gamow, G. 1935, *Ohio J. Sci.*, **35**, 406.
Gott, J. R., and Rees, M. J. 1975, *Astr. Ap.*, **45**, 365.
Gouda, N., and Sasaki, M., and Suto, Y. 1987, *Ap. J. (Letters)*, **321**, L1.
Guth, A. H., and Pi, S.-Y. 1982, *Phys. Rev. Letters*, **49**, 1110.
Harrison, E. R. 1970, *Phys. Rev.*, **1**, 2726.
Hawking, S. W. 1982, *Phys. Letters*, **115B**, 295.
Hogan, C. J. 1980, *M.N.R.A.S.*, **192**, 891.
———. 1984, *Ap. J. (Letters)*, **284**, L1.
Hogan, C. J., and Bond, J. R. 1988, in *Proc. NATO Advanced Study Institute, The Post-Recombination Universe*, ed. A. N. Lasenby and N. Kaiser (Dordrecht: Kluwer), p. 141.
Hogan, C. J., and White, S. D. M. 1986, *Nature*, **321**, 575.
Ikeuchi, S. 1981, *Pub. Astr. Soc. Japan*, **33**, 211.
Ikeuchi, S., Tomisaka, K., and Ostriker, J. P. 1983, *Ap. J.*, **265**, 583.
Kaiser, N. 1984*a*, *Ap. J.*, **282**, 374.
———. 1984*b*, *Ap. J. (Letters)*, **284**, L9.
———. 1986, *M.N.R.A.S.*, **222**, 323.
Kirschner, R. P., Oemler, A., Schechter, P. L., and Shectman, S. 1981, *Ap. J. (Letters)*, **248**, L57.
Klypin, A. A., Sazhin, M. V., Strukov, I. A., and Skulachev, D. P. 1987, *Soviet Astr. Letters*, **13**, 104.
Kodama, H., and Sasaki, M. 1986, *Internat. J. Mod. Phys.*, **A1**, 265.
Lake, G., and Partridge, R. B. 1980, *Ap. J.*, **237**, 378.
Lasenby, A. N. 1981, Ph.D. thesis, University of Manchester.
———. 1988, in *Large-Scale Motions in the Universe*, ed. V. C. Rubin, G. V. Coyne (Princeton: Princeton University Press), p. 278.
Lasenby, A. N., and Davies, R. D. 1983, *M.N.R.A.S.*, **203**, 1137.
Lawrence, C. R., Readhead, A. C. S., and Myers, S. 1988, in *Proc. NATO Advanced Study Institute, The Post-Recombination Universe*, ed. A. N. Lasenby and N. Kaiser (Dordrecht: Kluwer), p. 173.
Lehmann, E. L. 1986, *Testing Statistical Hypotheses* (New York: Wiley-Interscience).
Lemaître, G. 1927, *Ann. Soc. Sci. Brux.*, **A47**, 49.
Lifshitz, E. M. 1946, *Zh. Eksper. Teoret. Fiz.*, **16**, 587.
Martin, H. M., and Partridge, R. B. 1988, *Ap. J.*, **324**, 794.
Matsumoto, T., Hayakawa, S., Matsuo, H., Murakami, H., Sato, S., Lange, A. E., and Richards, P. L. 1988, *Ap. J.*, **329**, 567.
Moore, C. R. 1980, *IEEE Trans*, **MTT-28**, 149.
Moore, C. R., and Clauss, R. C. 1978, *IEEE Trans*, **MTT-27**, 249.
Ostriker, J. P., and Cowie, L. L. 1981, *Ap. J. (Letters)*, **243**, L127.
Ostriker, J. P., and Thompson, C. 1987, *Ap. J. (Letters)*, **323**, L97.
Ostriker, J. P., Thompson, C., and Witten, E. 1986, *Phys. Letters*, **B180**, 231.
Ostriker, J. P., and Vishniac, E. T. 1986, *Ap. J. (Letters)*, **306**, L51.
Parijskij, Yu. N. 1973*a*, *Ap. J. (Letters)*, **180**, L47.
———. 1973*b*, *Soviet. Astr.*, **17**, 291.
Parijskij, Yu. N., Berlin, A. B., and Vitkovskij, V. V. 1987, *Soob Spec. Astrofiz. Obs.*, **53**, 99.
Parijskij, Yu. N., Petrov, Z. E., and Cherkov, L. N. 1977, *Soviet. Astr. Letters*, **3**, 263.
Partridge, R. B. 1980*a*, *Phys. Scripta*, **21**, 624.
———. 1980*b*, *Ap. J.*, **235**, 681.
———. 1983, in *Proc. NATO Advanced Study Institute, The Origin and Evolution of Galaxies*, ed. B. L. T. Jones and J. E. Jones (Dordrecht: Reidel), p. 121.
———. 1989, in *Proc. 3d ESO/Cern Symposium, Astronomy, Cosmology, and Fundamental Physics*, ed. M. Caffo, R. Fanti, G. Giacomelli, and A. Renzini (Dordrecht: Kluwer), p. 105.
Pauliny-Toth, I. I. K., Witzel, A., Preuss, E., Baldwin, J. E., and Hills, R. E. 1978, *Astr. Ap. Suppl.*, **34**, 253.
Peebles, P. J. E. 1968, *Ap. J.*, **153**, 1.
———. 1974, *Ap. J. (Letters)*, **189**, L51.
———. 1987*a*, *Ap. J. (Letters)*, **315**, L73.
———. 1987*b*, *Ap. J.*, **317**, 576.
Peebles, P. J. E., and Yu, J. T. 1970, *Ap. J.*, **162**, 815.
Rees, M. J. 1972, *Phys. Rev. Letters*, **28**, 1669.
Rephaeli, Y. 1981, *Ap. J.*, **245**, 351.
Rubin, V. C., Thonnard, N., Ford, W. K., and Roberts, M. S. 1976, *A.J.*, **81**, 719.
Scaramella, R., and Vittorio, N. 1988, *Ap. J. (Letters)*, **331**, L53.
Scherrer, R. J. 1987, *Ap. J.*, **320**, 1.
Scheuer, P. A. G. 1957, *Proc. Camb. Phil. Soc.*, **53**, 764.
Silk, J. 1986, in *Inner Space/Outer Space*, ed. E. W. Kolb, M. S. Turner, K. Olive, D. Seckel, and D. Lindley (Chicago: University of Chicago Press), p. 126.
Silk, J., and Vittorio, N. 1987, *Ap. J.*, **317**, 564.
Silk, J., and Wilson, M. L. 1980, *Phys. Scripta*, **21**, 708.
Starobinsky, A. A. 1982, *Phys. Letters*, **117B**, 175.
Stebbins, A. 1988, *Ap. J.*, **327**, 584.
Steidel, C. C., and Sargent, W. L. W. 1987, *Ap. J. (Letters)*, **318**, L11.
Sunyaev, R. A. 1977, *Pis'ma Astr. Zh.*, **3**, 491.
———. 1978, *Soviet Astr. Letters*, **3**, 268.
Sunyaev, R. A., and Zel'dovich, Ya. B. 1972, *Astr. Ap.*, **20**, 189.
Uson, J. M., and Wilkinson, D. T. 1984*a*, *Ap. J. (Letters)*, **277**, L1.
———. 1984*b*, *Ap. J.*, **283**, 471.
———. 1984*c*, *Nature*, **312**, 427.
Vishniac, E. T. 1987, *Ap. J.*, **322**, 597.
Vishniac, E. T., and Ostriker, J. P. 1985, in *Proc. 3d Rome Meeting on Astrophysics, The Cosmic Background Radiation and Fundamental Physics*, ed. F. Melchiorri (Bologna: Italian Physical Society), p. 137.
Vittorio, N., and Silk, J. 1984, *Ap. J. (Letters)*, **285**, L39.
White, S. D. M., and Rees, M. J. 1978, *M.N.R.A.S.*, **183**, 341.
Wilson, M. L. 1983, *Ap. J.*, **273**, 2.
Wilson, M. L., and Silk, 1981, *Ap. J.*, **243**, 14.
Zel'dovich, Ya. B. 1967, *Soviet Phys. Usp.*, **9**, 602.
———. 1972, *M.N.R.A.S.*, **160**, 1P.
———. 1980, *M.N.R.A.S.*, **192**, 663.

H. E. HARDEBECK, C. R. LAWRENCE, S. T. MYERS, A. C. S. READHEAD, and W. L. W. SARGENT: California Institute of Technology, Mail Code 105-24, Pasadena, CA 91125

A. T. MOFFET: Deceased