

VLBA Observing Modes

JONATHAN D. ROMNEY
National Radio Astronomy Observatory
Charlottesville, Virginia

1985 May 5

INTRODUCTION

Recent decisions have solidified the VLBA recorder and correlator specifications sufficiently that a nearly-exhaustive enumeration of possible observing modes is now feasible. I attempt to do so in this memorandum, to provide a reference for other VLBA groups (and for interested members of the community), to suggest refinements the recorder and correlator groups should consider, and to aid in the evaluation of options in the current specifications. I have tried to organize the material to facilitate its use for all these purposes.

Section I is an exposition of the relevant parameters and their specifications from both the recorder and correlator areas. In Section II, I derive the optimal choices for the system-determined parameters in the full range of user-specified cases. The full range is then reduced in Section III to a major subset called basic observing modes, which should include all seriously conceivable cases. For these basic modes, Section IV presents the demands on the Array's resources. I propose standard observing modes for the most common configurations in Section V. Section VI concludes by suggesting areas for further discussion on rationalization of specifications and implementation of optional features.

I. PARAMETERS AND SPECIFICATIONS

For the purposes of this memorandum, I will consider a *parameter* to refer to some quantifiable aspect of the recorder or correlator system; a *specification* is an established statement of the permissible value(s) of the associated parameter; and a set of particular values of all parameters then defines an observing *mode*. Note that parameters may be (indeed, in our case generally are) mutually dependent. This implies interactions — and potential inconsistencies — between the associated specifications.

The following list enumerates all relevant parameters for which we have established specifications, and introduces nomenclature for subsequent reference. The specifications have been compiled from the plans of both the recorder and correlator groups.

It is useful to separate parameters into those designated by the user according to his or her scientific goals, and those whose values can be selected by the Array's control system for optimal performance.

USER-DESIGNATED PARAMETERS

These four parameters are the only aspects of the VLBA's observing mode of direct relevance to the user, who chooses their values to achieve the required sensitivity and spectral resolution. A large range of combinations exists, clearly, subject only to a limit imposed by the capacity of the VLBA recording system. Section V proposes optimal choices for standard observing modes.

Number of Channels: number of baseband channels recorded.

$$N_c = \{32, 16, 8, 4, 2, 1\} \text{ channels}$$

This specification represents a non-linear combination of the recorder and correlator specifications. Any number of channels up to 32 can be recorded; the correlator can process at most 16 channels simultaneously, and can deal usefully only with the binary submultiples of 16 which result from its tradeoff between number of channels and spectral resolution. My compromise includes the full 32-channel recorder capacity which may be useful in some cases even though it requires (at least) two processing passes, but I have restricted the generality of the recorder system to the binary submultiples which arise naturally in the correlator.

Channel Bandwidth: width of the narrowband filter applied at baseband.

$$B = \{8, 4, 2, 1, .5, .25, .125, .0625\} \text{ MHz/chnl}$$

The bandwidth is the only parameter which is formally independent of all others. Different bandwidths can be selected in separate channels.

Channel Sample Rate: rate at which the filtered baseband signal is sampled.

$$f_r = \{16, 8, 4, 2, 1, .5, .25, .125\} \text{ Msmpl/sec/chnl}$$

All recorded channels must be sampled at the same rate. Although the specified sample rates provide, obviously, Nyquist sampling for all possible channel bandwidths, VLBA specifications make the choice of f_r completely independent of B . Throughout this memorandum, however, I will assume sampling *at least* at the Nyquist rate (*i.e.*, undersampling will be forbidden). It will be seen below that the lower end of the specification range, $f_r \leq 1$ Msmpl/sec, cannot actually be used.

Sample Precision: number of recorded bits per sample.

$$\ell = \{1, 2\} \text{ bit/smpl}$$

The 2 bit/smpl case, it has now been decided, corresponds to 4-level quantization of the sampled signal. (This fact is irrelevant to most of the present discussion, but will have a powerful influence in the observer's choice of modes because it determines the sensitivity achieved.) I will use the terms *1-bit* and *2-level*, or *2-bit* and *4-level*, interchangeably according to context.

SYSTEM-DETERMINED PARAMETERS

The remaining parameters can be configured by the VLBA control system to optimize the Array's performance, as discussed in Section II. The consequent demands on the Array's primary resources — magnetic tape and processing time — are detailed in Section IV. These results may be of more than academic interest to the observer, since they will influence the favor with which his observing proposal will be received.

Number of Tracks: number of tape tracks recorded (on one or more recorders).

$$N_t = \{32, 16, 8, 4, 2, 1; 64, (128)\} \text{ tracks}$$

The limit of 32 simultaneously recorded tracks applies to a single recorder; we are planning to operate at least two recorders at each station. I have restricted the generality of the official specification to binary multiples and submultiples of 32 in parallel with the corresponding restriction in N_c .

Track Fanout: number of recorded tracks per channel bitstream.

$$\xi = \{1, 2, 4\} \text{ (trk/chnl) / (bit/smp)}$$

The term “fanout” applies to recording, of course; a corresponding fan-in occurs at playback. The complicated units arise because the fanout is applied to each sample bit in 2 bit/smp quantization. This quirk of the specifications in fact greatly simplifies the following analysis.

Record Rate: recorded bit rate per track.

$$g_r = \{8, 4, 2\} \text{ Mbit/sec/trk}$$

I have chosen to specify “tape speed” in the units fundamental to the discussion of observing modes. For a given longitudinal bit density the record rate is proportional to tape speed; under current VLBA recorder specifications the rates above correspond to *nominal* speeds of 240, 120, and 60 ips. (The actual record rates and tape speeds will be inflated by more than 12% to allow for parity bits and framing blocks.) In specific textual references it’s less awkward to refer to these rates as “full”, “half”, and “quarter” speed; these designations are unfortunately at variance with those current at Haystack, but appear more suitable to this discussion.

Speedup Factor: ratio of playback to record rate.

$$\alpha = \{(\frac{1}{2}), 1, 2, (4)\}$$

The specification refers to support in the correlator, where the speedup has the effect of *lengthening* the observe-time intervals corresponding to fixed intervals of playback time, so that geometric calculations in the correlator become less accurate. The $(\frac{1}{2})$ specification requires the optional half-speed playback described in the next paragraph; the entry (4) represents a partially-supported value, permissible if bandwidth $B \leq 4$ MHz and observing frequency $\nu \leq 22$ GHz.

Playback Rate: reproduced bit rate per track.

$$g_p = \{8, (4)\} \text{ Mbit/sec/trk}$$

All the comments under “record rate” above apply here as well, except that the half-speed playback is currently an option which may or may not be implemented, and no quarter-speed playback is planned. The intent of the half-speed playback option is to reduce the speedup factor and thus the length of observe-time intervals which correspond to the fixed playback-time intervals specified for certain correlator functions. In particular, it is a prerequisite for the $\alpha = \frac{1}{2}$ slowdown. If this option is available, however, it may be useful in other connections too.

Playback Sample Rate: rate at which samples are delivered to the correlator.

$$f_p = \{16, 8, 4, 2, 1, .5, .25, .125, \dots\} \text{ Msmp/sec/chnl}$$

This parameter refers to the “logical” sample rate; the correlator will probably receive actual samples at a fixed 16 Msmp/sec/chnl rate, and may decimate these according to its internal shift clock. The specification is (presumably) identical to that for f_r ; the essential feature for later discussion, however, is the 16 Msmp/sec/chnl maximum.

DEPENDENCIES

Only six of the preceding parameters are independent, *i.e.*, only two of the system-determined parameters are independent once the user-designated parameters are known. The mutual dependencies are given by the equations

$$\begin{aligned} f_r &= \xi g_r & \alpha &= \frac{g_p}{g_r} = \frac{f_p}{f_r}, \\ N_t &= \xi \ell N_c \end{aligned} \tag{1}$$

involving all parameters except B , which is (formally) independent. (A close, but not unique, coupling between B and f_r will be introduced in Section III.)

II. OPTIMIZATION OF PARAMETERS

It will be clear from the previous section that the set of parameters for which we have established specifications is highly mutually dependent, so that the ranges of permissible values for the system-determined parameters are restricted by a variety of interacting specifications. Within these ranges the parameter values must be optimized according to as yet unspecified criteria. These two topics are treated in the present section.

PARAMETER RANGES

For the purpose of elucidating the restrictions on the system-determined parameters, it is most convenient to regard the fanout ξ and speedup α as the independent variables. Then each of the equations (1) displays a relatively independent dimension of the restrictions. Although these restrictions are really quite obvious, it's also clear that some have never been realized before; I hope it will not be too tedious to point them out explicitly in the following brief paragraphs.

Channel Sample Rate. Given the track fanout ξ , the user-designated channel sample rate f_r determines the record rate g_r . From the specifications of these two parameters, nine combinations of ξ and g_r would appear to exist, covering five sample rates in octave steps from 2 to 32 Msmp/sec/chnl. However, the greatest of these rates violates the f_r specification and this combination must always be excluded. Conversely, there are no combinations which can realize $f_r \leq 1$ Msmp/sec, so that these specified sample rates *cannot be recorded at all*; this in turn implies that narrow observing bandwidths $B < 1$ MHz *must* be oversampled by (at least) a factor of $1 \text{ MHz}/B$.

Playback Sample Rate. Additional restrictions arise when the effect of the speedup factor α is considered. To avoid violating the maximum f_p specification, clearly we may only invoke $\alpha = 2$ for $f_r \leq 8$ Msmp/sec/chnl, and $\alpha = 4$ for $f_r \leq 4$ Msmp/sec/chnl. These restrictions, incidentally, are more conservative by an octave (at least — more for oversampling) than the official limitations based on bandwidth which arise from the accuracy of delay and phase tracking in the correlator.

Number of Tracks. Finally, the limit of 32 simultaneously recorded tracks further constrains the permissible combinations of system-determined parameter values. This is not a hard limit, since we are planning to have at least two recorders available at each station. However, using both recorders in parallel impairs somewhat the reliability of the Array, so I have chosen to enforce this limit, except for observations that are not otherwise possible. Generally this imposes a restriction on the usable fanout factors ξ when the product $\ell N_c \geq 16$ in equations (1); the exceptions are extreme wideband observations, which require a large number of high-sample-rate channels.

Table 1 summarizes the relationships between the user-specified channel sample rate f_r , the combinations of ξ with g_r , and the corresponding combinations of α and f_p for both the official and the optional playback rates g_p . All possible combinations appear in the table, with the illegal cases shown lined through in the columns where the specifications are violated.

TABLE 1
PARAMETER RANGES

Channel Sample Rate f_r (Msmp/sec /chnl)	Track Fanout ξ (trk/chnl) /(bit/smp)	Record Rate g_r (Mbit/sec /trk)	Playback Rate g_p (Mbit/sec/chnl)			
			8		4	
			Speedup Factor α	Playback Sample Rate f_p (Msmp/sec /chnl)	Speedup Factor α	Playback Sample Rate f_p (Msmp/sec /chnl)
32	4	8	1	32	$\frac{1}{2}$	16
16	4	4	2	32	1	16
	2	8	1	16	$\frac{1}{2}$	8
8	4	2	4	32	2	16
	2	4	2	16	1	8
	1	8	1	8	$\frac{1}{2}$	4
4	2	2	4	16	2	8
	1	4	2	8	1	4
2	1	2	4	8	2	4

One interesting fact which becomes evident in Table 1 is that the maximum fanout $\xi = 4$ *cannot be used at all with full-speed playback*. The reason is obvious but somewhat circuitous: this fanout induces a record rate sufficiently slow that the speedup for full-speed playback violates the maximum playback sample rate specification.

OPTIMIZATION STRATEGY

A choice of parameter combinations exists for at least some of the channel sample rates in Table 1, especially if the optional half-speed playback is available. How should we define the optimum combination? I have adopted the following primary strategy: *choose the slowest possible record rate* (and thus the largest track fanout), *under the constraint imposed by the limited number of tracks*. In some cases the “slowest possible” record rate may depend on the availability of the optional half-speed playback. This strategy minimizes the number of tape passes (and thus tape and head wear), and maximizes the speedup factor (allowing multi-pass processing of this or some other observation).

In some specialized observations, the opposite strategy may be warranted: *minimize* the speedup factor by choosing the fastest possible record rate (and, if available, the half-speed playback). This “widefield mode” ensures the shortest observe-time intervals corresponding to the fixed playback-time intervals for correlator readout and delay and phase computation; besides its primary purpose of expanding the observable field of view, it may also be valuable in high-frequency observations to preserve phase-tracking accuracy. If the optional half-speed playback is implemented, this mode actually yields a *slowdown*, $\alpha = \frac{1}{2}$, for $f_r = \{16, 8\}$ Msamp/sec/chnl. At narrow bandwidths, however, this induces extreme oversampling — a factor of 4 beyond that required to reach the slowest record rate.

Tables 2a–b present the optimum parameter values under the primary strategy, at opposite extremes of conciseness/completeness. Table 2a shows only ξ and α , for the four permissible channel sample rates f_r , with the influence of the N_t limit expressed in terms of the product ℓN_c ; “widefield mode” (for any ℓN_c) is shown as an additional case. The remaining parameters can be obtained via equations (1), of course. It’s often convenient to have all the parameter values explicitly visible at once, so Table 2b presents ξ , g_r , N_t , and α in a foursquare element for all combinations of f_r , ℓ , and N_c (but not for widefield mode). All units are the same as in Table 1.

TABLE 2a
CONCISE OPTIMIZED PARAMETERS

$\xi : \alpha$				
f_r	$\ell N_c \geq 32$	$\ell N_c = 16$	$\ell N_c \leq 8$	Widefield
16	2 : 1	2 : 1	4 : 1	2 : $\frac{1}{2}$
8	1 : 1	2 : 2	4 : 2	1 : $\frac{1}{2}$
4	1 : 2		2 : 4	1 : 1
2	1 : 4		1 : 4	1 : 2

TABLE 2b
COMPLEAT OPTIMIZED PARAMETERS

		$\xi = \text{track fanout}$ $N_t = \text{number of tracks}$	$\frac{\xi g_r}{N_t \alpha}$	$g_r = \text{record rate}$ $\alpha = \text{speedup factor}$			
f_r	ℓ	$N_c = 32$	16	8	4	2	1
16	1	$\frac{2 8}{64 1}$	$\frac{2 8}{32 1}$	$\frac{4 4}{32 1}$	$\frac{4 4}{16 1}$	$\frac{4 4}{8 1}$	$\frac{4 4}{4 1}$
	2	$\frac{2 8}{128 1}$	$\frac{2 8}{64 1}$	$\frac{2 8}{32 1}$	$\frac{4 4}{32 1}$	$\frac{4 4}{16 1}$	$\frac{4 4}{8 1}$
8	1	$\frac{1 8}{32 1}$	$\frac{2 4}{32 2}$	$\frac{4 2}{32 2}$	$\frac{4 2}{16 2}$	$\frac{4 2}{8 2}$	$\frac{4 2}{4 2}$
	2	$\frac{1 8}{64 1}$	$\frac{1 8}{32 1}$	$\frac{2 4}{32 2}$	$\frac{4 2}{32 2}$	$\frac{4 2}{16 2}$	$\frac{4 2}{8 2}$
4	1	$\frac{1 4}{32 2}$	$\frac{2 2}{32 4}$	$\frac{2 2}{16 4}$	$\frac{2 2}{8 4}$	$\frac{2 2}{4 4}$	$\frac{2 2}{2 4}$
	2	$\frac{1 4}{64 2}$	$\frac{1 4}{32 2}$	$\frac{2 2}{32 4}$	$\frac{2 2}{16 4}$	$\frac{2 2}{8 4}$	$\frac{2 2}{4 4}$
2	1	$\frac{1 2}{32 4}$	$\frac{1 2}{16 4}$	$\frac{1 2}{8 4}$	$\frac{1 2}{4 4}$	$\frac{1 2}{2 4}$	$\frac{1 2}{1 4}$
	2	$\frac{1 2}{64 4}$	$\frac{1 2}{32 4}$	$\frac{1 2}{16 4}$	$\frac{1 2}{8 4}$	$\frac{1 2}{4 4}$	$\frac{1 2}{2 4}$

Both tables assume that the optional half-speed playback is implemented and take advantage of this feature where appropriate. If only full-speed playback is available, wide-field mode becomes identical to the case $\ell N_c \geq 32$, and the non-widefield configurations identified by the occurrence of $\xi = 4$ must be modified according to

$$\begin{aligned} \xi' &= 2 & g_r' &= 2g_r \\ N_t' &= N_t/2 & \alpha' &= \alpha \end{aligned} \quad (2)$$

III. BASIC OBSERVING MODES

Further analysis is obstructed by the considerable volume of parameter space available to the user under the formal specifications. Much of this volume in fact represents modes of negligible utility. Therefore I define in this section a major subset of observing modes, called *basic modes*, which reduces the mode multiplicity to a manageable level while preserving all the essential dimensions of flexibility. This reduction is accomplished by reconsidering the coupling between the user-designated bandwidth B and the channel sample rate f_r .

OVERSAMPLING

Formally there is *no* coupling between B and f_r . I have assumed that *undersampling* will not be permitted, and we saw in Section II that for $B < 1$ MHz *oversampling* will be unavoidable. Nevertheless these modifications still allow more flexibility in oversampling than is ever likely to be useful. To determine what range of oversampling we really need to support we must consider its benefits in more detail. Oversampling is attractive to the user because it enhances sensitivity in two distinct ways, indicated in Table 3.

TABLE 3
SENSITIVITY ENHANCEMENTS FROM OVERSAMPLING
(Columns normalized to unity at infinite oversampling factor)

Oversampling Factor	Extra bits correlated		Delay tracking refinement		
	2-level Sampling	4-level Sampling	Band Edge	Quartile Points	Mean Over Band
1	.797	.911	.900	.974	.966
2	.927	.965	.974	.994	.991
4	.977	.985	.994	.998	.998
8	.992	.993	.998	1.000	.999
16	.996	.997	1.000	1.000	1.000

One enhancement, common to all digital correlators, results from correlating the over-sampled bits, with some improvement in signal-to-noise ratio. This effect was evaluated by Bowers and Klingler (A&A Suppl. Series, Vol. 15, p. 373, 1974), and their results are extracted in the table; the enhancement is stronger for 2-level than for 4-level sampling, but is present in both. The VLBA correlator will be able to correlate these oversampled bits, but current specifications do not allow for extra lag steps between correlated lags to retain the full spectral resolution, except as an option for oversampling by a factor of 2.

A second sensitivity enhancement is particular to digital interferometers. The finer time resolution in the bit streams allows correspondingly finer tracking of the interferometer delay, which in turn reduces the frequency-dependent phase-tracking errors and the consequent decorrelation away from the phase-tracking reference frequency. Under the usual assumptions (frequent delay "bit-shifts" within an integration period, and phase tracking at the RF band center), I have tabulated the decorrelation at the band edges and quartile points, and a mean decorrelation over the band.

Table 3 shows clearly that oversampling by a factor of 4 suffices to achieve virtually the entire sensitivity enhancement in all cases. The factor of 2 oversampling is probably also worth retaining because it is less expensive and may be fully supported by the correlator. I propose then to reduce the oversampling dimension to one primary state, *standard sampling*, and two secondary states, *double* and *quadruple sampling*. These names are meant,

obviously, to imply oversampling by factors of 1, 2, and 4, respectively, but since only four channel sample rates are available the desired oversampling factors can actually be achieved only in a narrow range of bandwidths. Because it frequently forces a gross degree of oversampling at the sample rates necessary to minimize the speedup factor, the table also illustrates "widefield mode" with $\alpha = \frac{1}{2}$ as an additional (tertiary) sampling state.

These four cases are summarized in Table 4, where the linkages between bandwidth B and channel sample rate f_r are shown explicitly. Under standard sampling and widefield mode I indicate as well the *forced* oversampling factors in square brackets; the entries for the secondary sampling states appear only where the nominal oversampling factor is actually realizable *and* different from that for standard sampling.

TABLE 4

SAMPLING STATES

f_r (Msamp/sec/chnl) [oversampling]

B (MHz/chnl)	Sampling State			
	Standard	Double	Quadruple	Widefield
8.	16	—	—	16
4.	8	16	—	8
2.	4	8	16	8 [2]
1.	2	4	8	8 [4]
.5	2 [2]	—	4	8 [8]
.25	2 [4]	—	—	8 [16]
.125	2 [8]	—	—	8 [32]
.0625	2 [16]	—	—	8 [64]

SAMPLE PRECISION

One further simplification is accomplished by considering carefully the linkage between the sampling states and the sample precision, *i.e.*, the number of quantization levels. In tape-limited observations (primarily wideband continuum measurements) the two cases of 1-bit sampling at a given bandwidth, and 2-bit sampling at half the same bandwidth, achieve virtually identical sensitivities while recording the same number of bits. This decision then is likely to be dominated by other considerations, such as the playback speedup factor, volume of archive data generated, or compatibility with non-VLBA systems. Oversampling, however, is clearly undesirable because it reduces the bandwidth for a given number of recorded bits, with less than an offsetting gain in sensitivity. Conversely, in band-limited cases (*e.g.*, spectroscopic, pulsar, or interference-prone low-frequency observations), sensitivity gains may be achieved both by increasing the sample precision from 1 to 2 bits, and by oversampling. The former is a far more effective use of the extra recorded bits; thus *there is no reason to support the 1-bit oversampled combinations.*

Narrow bandwidths in general, and especially in “widefield mode”, represent a case where the oversampling is inadvertent and not necessarily motivated by a requirement for ultimate sensitivity. In this case the preceding dictum would impose 2-bit sampling where it’s not really desirable. In general, I regard this unpleasant wrinkle as a reasonable price for a major simplification, since the penalty is not severe for ordinary narrow-band observations. But in widefield mode the required high sample rate makes the 2-bit case quite expensive in recorded bits. Nevertheless, I can see some application for both 1- and 2-bit versions, and I suppose if we are serious about this option we should consider both.

DEFINITION OF BASIC MODES

On the basis of the preceding discussion, I propose to define the following *basic observing modes* for the VLBA. There are two basic sampling dimensions, the sampling *state* and sample *precision*, but only six combinations are included in the basic modes:

- Standard 1-bit.** Closest to current VLBI practice; yields very highest sensitivity in tape-limited observations.
- Standard 2-bit.** Slightly ($\sim 3\%$) less sensitive than 1-bit tape-limited case but may reduce archive volume; major sensitivity gain for band-limited observations.
- Double 2-bit.** Enhances band-limited sensitivity; spectral resolution in correlator reduced by factor of 2, or may (optionally) be supported at full resolution.
- Quadruple 2-bit.** Ultimate sensitivity enhancement; spectral resolution reduced by factor of 4 or (optionally) 2.
- Widefield 1-bit.** Provides minimum possible speedup factor; suitable for widefield or high-frequency continuum observations.
- Widefield 2-bit.** Uses even more tape than 1-bit version; may be necessary for high-sensitivity spectroscopy.

The two remaining dimensions of the basic mode space are the bandwidth B and number of channels N_c . In each sampling combination, the sampling state determines a specific channel sample rate f_r from B as shown in Table 4, and the sample precision ℓ combines with N_c to determine all other parameters according to the results of Section II.

Perhaps it’s worth repeating that *these basic observing modes should suffice for any seriously conceivable observation*. Only the following cases have been excluded: undersampling, unreasonable elective oversampling (*i.e.*, by more than a factor of 4), oversampling combined with 2-level quantization, and widefield cases with intermediate speedups.

IV. ARRAY RESOURCE REQUIREMENTS

The discussion of the previous section makes possible an explicit analysis of the impact of various observing modes on the Array’s primary resources: magnetic tape and processing time. (Allocation of observing time, *the* primary resource, will in turn be affected by these resource demands, but this is not likely to be a quantifiable relationship. Demand for post-processing capacity, another major resource, will also be heavily mode-dependent, but that analysis is beyond the scope of this memorandum.)

Two important operational details for both tape consumption and processing time are the following *multipass factors*. K_t is the number of recorders used simultaneously at each Array element; these tapes must be correlated separately because the playback system will have only one unit available per station (unless the observations are restricted to 10 or fewer stations). And K_c reflects the correlator's limitation to 16 channels, so that 32-channel observations must be correlated in two passes. The factors are simply

$$K_t = \left\lceil \frac{N_t}{32 \text{ trk}} \right\rceil \quad \text{and} \quad K_c = \left\lceil \frac{N_c}{16 \text{ chnl}} \right\rceil. \quad (3)$$

(I have chosen to assume no particular limit on K_t , but rather let the extreme cases speak for themselves.)

To quantify magnetic tape consumption some assumptions will be necessary about the eventual capacity of the VLBA record system. Although no final decision has been reached, recent information suggests that it will be feasible to record 32 tracks at 4 Mbit/sec/trk for 12 hours on a single reel. (For the curious, this assumes 28 passes on an 18000-foot reel at a tape speed of 135 ips, but these numbers are not otherwise relevant.) As a general parametrization I will define the overall *tape capacity* as

$$T = \frac{32 \text{ trk}}{N_t} \cdot \frac{4 \text{ Mbit/sec/trk}}{g_r} \cdot 12 \text{ hr/tape}. \quad (4)$$

For many purposes, however, it's important to know K_t as well; for $K_t > 1$, I will follow the convention of writing T as the unreduced fraction $\frac{TK_t}{K_t}$.

Processing time is best represented by the overall processing speedup factor. This is frequently equal to the playback speedup factor α , but in extreme cases must be reduced to reflect forced multi-pass processing. Accordingly, the *processing speedup* factor is

$$P = \alpha / \max\{K_t, K_c\}. \quad (5)$$

I tabulate below T and P for all the basic observing modes defined in the previous section. Table 5a includes the most important modes, 1- and 2-bit standard sampling; the extreme high-data-rate cases towards the upper left corner are unrealistic for sustained observations under current operational plans, but may be valuable in short bursts at high frequencies. Table 5b presents the secondary 2-bit oversampling modes. Note that while the extreme wideband oversampled modes are omitted because they cannot be implemented, the extreme narrowband cases do not appear because standard sampling already involves a forced oversampling. Finally, Table 5c shows the cost in T and P for the 1- and 2-bit "widefield modes" with $\alpha = \frac{1}{2}$.

The entries in Tables 5a–b are *not* conditional upon implementation of the optional half-speed playback. Equations (2) and (4) show that T is unchanged if only full-speed playback is assumed; this is obvious enough since the same total number of bits must be recorded. Under the primary optimization strategy of Section II, the N_t limit ensures that half-speed full-fanout recording is only considered when $\ell N_c \leq 8$ — as can be seen in Table 2a — so that K_t in equation (3) is unaffected by the halving of N_t , and equations (2) and (5) show that P too is unchanged. In Table 5c, however, P must be doubled in all entries if half-speed playback is not implemented.

TABLE 5
ARRAY RESOURCE REQUIREMENTS

$T:P$

(a) Standard Sampling

B	ℓ	$N_c = 32$	16	8	4	2	1
8.	1	$\frac{6}{2}:\frac{1}{2}$	6:1	12:1	24:1	48:1	96:1
	2	$\frac{6}{4}:\frac{1}{4}$	$\frac{6}{2}:\frac{1}{2}$	6:1	12:1	24:1	48:1
4.	1	$6:\frac{1}{2}$	12:2	24:2	48:2	96:2	192:2
	2	$\frac{6}{2}:\frac{1}{2}$	6:1	12:2	24:2	48:2	96:2
2.	1	12:1	24:4	48:4	96:4	192:4	384:4
	2	$\frac{12}{2}:1$	12:2	24:4	48:4	96:4	192:4
$\leq 1.$	1	24:2	48:4	96:4	192:4	384:4	768:4
	2	$\frac{24}{2}:2$	24:4	48:4	96:4	192:4	384:4

(b) Double and Quadruple Sampling ($\ell = 2$ only)

B	Sampling	$N_c = 32$	16	8	4	2	1
4.	Double	$\frac{6}{4}:\frac{1}{4}$	$\frac{6}{2}:\frac{1}{2}$	6:1	12:1	24:1	48:1
2.	Double	$\frac{6}{2}:\frac{1}{2}$	6:1	12:2	24:2	48:2	96:2
	Quadruple	$\frac{6}{4}:\frac{1}{4}$	$\frac{6}{2}:\frac{1}{2}$	6:1	12:1	24:1	48:1
1.	Double	$\frac{12}{2}:1$	12:2	24:4	48:4	96:4	192:4
	Quadruple	$\frac{6}{2}:\frac{1}{2}$	6:1	12:2	24:2	48:2	96:2
.5	Quadruple	$\frac{12}{2}:1$	12:2	24:4	48:4	96:4	192:4

(c) Widefield Mode ($\alpha = \frac{1}{2}$)

B	ℓ	$N_c = 32$	16	8	4	2	1
8.	1	$\frac{6}{2}:\frac{1}{4}$	$6:\frac{1}{2}$	$12:\frac{1}{2}$	$24:\frac{1}{2}$	$48:\frac{1}{2}$	$96:\frac{1}{2}$
	2	$\frac{6}{4}:\frac{1}{8}$	$\frac{6}{2}:\frac{1}{4}$	$6:\frac{1}{2}$	$12:\frac{1}{2}$	$24:\frac{1}{2}$	$48:\frac{1}{2}$
$\leq 4.$	1	$6:\frac{1}{4}$	$12:\frac{1}{2}$	$24:\frac{1}{2}$	$48:\frac{1}{2}$	$96:\frac{1}{2}$	$192:\frac{1}{2}$
	2	$\frac{6}{2}:\frac{1}{4}$	$6:\frac{1}{2}$	$12:\frac{1}{2}$	$24:\frac{1}{2}$	$48:\frac{1}{2}$	$96:\frac{1}{2}$

V. STANDARD OBSERVING MODES

The *standard observing modes* defined in this section are intended to provide straightforward choices for the casual VLBA user, and to suggest priorities (but not scope!) for design and implementation within the VLBA project. More experienced users may refer to the full set of *basic modes* defined in Section III and characterized in Section IV. In the following enumeration, the standard modes are grouped according to the astronomical function of the observation.

As has been the case throughout this memorandum, I do not distinguish between observations which obtain polarization information and those which do not. In “polarization” observations the N_c channels are assigned to $N_c/2$ pairs in orthogonal polarizations at the same RF frequency within each pair; these then yield $N_c/2$ sets of complete Stokes parameters. This specific choice of polarization and frequency has no particular impact on the recording and playback systems at all. But the polarization cross-products reduce the correlator’s channel limit to $N_c \leq 8$ and the spectral resolution by a factor of 4. These effects are documented in VLBA Correlator Memo VC 041.

In defining the standard modes, frequent reference is necessary to the *fundamental VLBA recording specification* of 24 hours unattended operation, using 2 recorders.

CONTINUUM MODES

Continuum observations are characterized by the frequent necessity for maximal total observing bandwidths. They are thus generally “tape-limited”, and as discussed in Section III under “sample precision” both 1-bit and 2-bit sampling — without oversampling — must be considered. Two-bit sampling requires only half the bandwidth, and generates half as much correlated data, for a given sensitivity, and we can expect users to migrate to this mode eventually; however, 1-bit sampling offers the advantage of greater compatibility with the existing Mark III system, and will probably be used heavily in the early years of VLBA observations. Hence, I have defined standard modes for both cases.

Because the detailed recording specifications allow significantly higher data rates than the average rate which can be sustained under the fundamental specification, I have adopted the following supermarket-style nomenclature for the continuum modes.

Normal Continuum Modes. These modes span a total bandwidth of 64 MHz with 1-bit sampling (32 MHz with 2-bit), and allow sustained observing at the VLBA’s basic data rate, with a tape capacity of 12 hr/tape. Inspection of Table 5 shows that we must consider a trade-off here between the volume of correlated data (*i.e.*, the number of channels) and the processing speedup. Since these modes will likely accommodate the preponderance of VLBA observations, I have chosen the standard to minimize the bulk of archive data while processing without a speedup. Continuum observations requiring reprocessing will probably deviate from this standard to achieve the speedup.

Normal Continuum 1-bit: 8 8-MHz channels, 12 hr/tape, no speedup.

Normal Continuum 2-bit: 4 8-MHz channels, 12 hr/tape, no speedup.

Super Continuum Modes. With double the bandwidth of the corresponding “normal” cases, these modes reduce the tape capacity to 6 hr/tape and thus cannot be sustained under current operational plans. However, in mixed observations one could observe for, *e.g.*, 8 hours in one of the super modes and 16 hours in a low-resolution spectral mode (defined below) at a bandwidth of 1 MHz or narrower without violating the fundamental specification of 24 hours of unattended operation. At the data rates required for these modes, no processing speedup is possible, and no trade-off need be considered.

Super Continuum 1-bit: 16 8-MHz channels, 6 hr/tape, no speedup.

Super Continuum 2-bit: 8 8-MHz channels, 6 hr/tape, no speedup.

Ultra Continuum Modes. The VLBA’s maximum sensitivity is achieved in these modes, by recording simultaneously on two recorders at maximum rate. A typical combination with a low-data-rate mode would juxtapose 4.8 hours in an ultra mode with 19.2 hours of high-resolution spectral observations (at any bandwidth). A processing *slowdown* is the inevitable consequence of dual-recorder recording.

Ultra Continuum 1-bit: 32 8-MHz channels, 6 hr/2 tapes, $2\times$ *slowdown*.

Ultra Continuum 2-bit: 16 8-MHz channels, 6 hr/2 tapes, $2\times$ *slowdown*.

Widefield Continuum Modes. Following the principle that we should only push the limits of one system parameter at a time, I have defined the widefield modes at the “normal” continuum sensitivity. Since the slowdown is the essence of these modes, there is no trade-off in this case between the volume of correlated data and processing speed. However, a different trade-off arises from the “delay beam” restriction of the field of view, which will dominate the fringe-rate window at low frequencies (how low depends on the baselines used). To minimize this effect I have chosen a larger number of narrower channels than in the normal continuum case. The additional volume of correlated data will not be a major burden for these relatively infrequent observations.

Widefield Continuum 1-bit: 16 4-MHz channels, 12 hr/tape, $2\times$ *slowdown*.

Widefield Continuum 2-bit: 8 4-MHz channels, 12 hr/tape, $2\times$ *slowdown*.

SPECTROSCOPIC MODES

In contrast to continuum observations, spectroscopy is generally band-limited, and as discussed in Section III, both 2-bit sampling and oversampling yield significant sensitivity enhancements. On the other hand, with a bandwidth matched to the spectral lines, the number of channels (*i.e.*, number of lines) required is typically small. In these modes the Array’s throughput is limited by the number of correlator lags available, and the number of channels must be traded off against the desired spectral resolution (as specified in VLBA Correlator Memo VC 041). The demands on the recording system, however, are generally modest — even in the presence of fairly extreme oversampling.

I have defined both low- and high-resolution spectroscopic modes below, differing by a factor of 4. In both cases the number of channels is small enough that 2-bit sampling with maximal oversampling (*i.e.*, within the range available under the basic observing modes) can reasonably be accommodated.

High-Resolution Spectroscopic Modes (2-bit). Maximum spectral resolution for a given bandwidth is obtained when only *1 channel* is recorded and correlated: 512/256/128 spectral channels for 10/14/20 stations. The tape consumption in this case is at most a quarter of the fundamental VLBA recording capacity. The following modes are drawn from Tables 5a–b, as appropriate to achieve maximal oversampling.

- High-Resolution 8 MHz:** standard sampling, 48 hr/tape, no speedup.
- High-Resolution 4 MHz:** double sampling, 48 hr/tape, no speedup.
- High-Resolution 2 MHz:** quadruple sampling, 48 hr/tape, no speedup.
- High-Resolution 1 MHz:** quadruple sampling, 96 hr/tape, 2× speedup.
- High-Resolution 0.5 MHz:** quadruple sampling, 192 hr/tape, 4× speedup.
- High-Res. Narrowband:** $\geq 4\times$ oversampling, 384 hr/tape, 4× speedup.

Low-Resolution Spectroscopic Modes (2-bit). In some cases it may be valuable to observe several lines with reduced resolution. These modes yield 128/64/32 spectral channels (again for 10/14/20 stations) in each of *4 channels*, and parallel exactly the high-resolution modes in sampling and speedup, with 4 times the tape consumption (which, however, never exceeds the rate sustainable under the fundamental recording specification).

- Low-Resolution 8 MHz:** standard sampling, 12 hr/tape, no speedup.
- Low-Resolution 4 MHz:** double sampling, 12 hr/tape, no speedup.
- Low-Resolution 2 MHz:** quadruple sampling, 12 hr/tape, no speedup.
- Low-Resolution 1 MHz:** quadruple sampling, 24 hr/tape, 2× speedup.
- Low-Resolution 0.5 MHz:** quadruple sampling, 48 hr/tape, 4× speedup.
- Low-Res. Narrowband:** $\geq 4\times$ oversampling, 96 hr/tape, 4× speedup.

Widefield Spectroscopic Modes (2-bit). Again restricting extreme configurations to a single system parameter wherever possible, I have defined the widefield spectroscopic modes avoiding oversampling where possible. Since a minimum sample rate of 8 Msamp/sec/chnl is necessary to achieve the required slowdown in processing speed, however, in fact all bandwidths narrower than 4 MHz are subject to a forced oversampling. Both the high- and low-resolution cases are combined in the following list.

- Widefield High-Res. Wideband:** 1 channel, 8 MHz, 48 hr/tape, 2× *slowdown*.
- Widefield High-Res. Normal:** 1 channel, ≤ 4 MHz, 96 hr/tape, 2× *slowdown*.
- Widefield Low-Res. Wideband:** 4 channels, 8 MHz, 12 hr/tape, 2× *slowdown*.
- Widefield Low-Res. Normal:** 4 channels, ≤ 4 MHz, 24 hr/tape, 2× *slowdown*.

PULSAR MODES

The prime consideration in pulsar observations is the requirement for relatively narrow bands to facilitate de-dispersion of the pulses. A large number of such channels are then necessary to detect the weak pulses satisfactorily; 2-bit sampling contributes to this end as well. These modes were suggested by Alan Rogers in VLBA Memo 435 (although the internal details — track fanouts and record rates — must deviate from his particulars).

- Normal Pulsar:** 16 1-MHz channels, 2-bit, 24 hr/tape, 4× speedup.
- High-Sensitivity Pulsar:** 32 1-MHz channels, 2-bit, 24 hr/2 tapes, 2× speedup.

BANDWIDTH-SYNTHESIS MODES

These modes, intended primarily for geodetic and astrometric measurements, require a large number of channels for satisfactory synthesis. Sensitivity, however, is not a major consideration because strong sources are sufficiently dense on the sky. Thus I have specified 2 MHz bands observed with 1-bit standard sampling, in parallel with current Mark III practice, for the "normal" mode. The high-sensitivity case may be required for phase-referencing work where a nearby reference is essential. Both modes are compatible with Mark III Mode C observations, except in the precise number of channels (and, of course, in the details of track width, data format, *etc.*).

Normal BWS: 16 2-MHz channels, 1-bit, 24 hr/tape, 4× speedup.

High-Sensitivity BWS: 16 4-MHz channels, 1-bit, 12 hr/tape, 2× speedup.

VI. TOPICS FOR FURTHER DISCUSSION

This memorandum has emphasized primarily an expansion of current recorder and correlator specifications to examine their consequences, and as such its major results consist in large part of the tabulated parameters and mode definitions. These expanded results cannot sensibly be summarized, and I have not attempted a general summary. However, this analysis has revealed some anomalous connections and even inconsistencies among our specifications, and these I recapitulate in this section, with recommendations for changes. Furthermore, the implementation options, particularly the ubiquitous half-speed playback, have extensive implications for the mode definitions. I conclude by discussing these options and the Array capabilities they are designed to support, emphasizing their relation to the observing modes.

SPECIFICATION ANOMALIES

These anomalies were noted in the discussion of "parameter ranges" in Section II. They arise in the connections among the track fanout ξ , and the speedup factor α which relates the record rate g_r and playback rate g_p , and the channel sample rate f_r and playback sample rate f_p .

Track Fanout and Record Rate vs. Channel Sample Rate. The ranges specified for ξ and g_r are inconsistent with both ends of the f_r specification. Only four of the five unique products ξg_r are within the limit $f_r \leq 16 \text{ Msmp/sec/chnl}$; this inefficiency of the specifications is treated further in the next paragraph. More importantly, these combinations are restricted to $f_r \geq 2 \text{ Msmp/sec/chnl}$, so that the lower range of the f_r specification cannot be recorded. This implies a forced oversampling for narrowband observations, with a resulting unnecessary use of tape. Extending the range of achievable values of f_r downwards would require either implementing a track *fan-in* to reach fractional values of ξ , or recording at even slower rates than $g_r = 2 \text{ Mbit/sec/trk}$. As the latter does not appear to be feasible with acceptable tape positioning using the narrow-track VLBA heads, we must consider whether the cost of multiplexing to achieve the fan-in is worth the savings in tape consumption. Or, as some have argued, has the point of diminishing returns already been reached? In either case, however, the f_r specification should include only those values actually available.

Track Fanout and Playback Rate vs. Playback Sample Rate. The maximum fanout $\xi = 4$ is not required to achieve any allowable $f_r \leq 16$ Msmp/sec/chnl. And $\xi = 4$ combined with *any* allowable g_r and the *standard* $g_p = 8$ Mbit/sec/trk, imply an excessive $f_p > 16$ Msmp/sec/chnl. Only with the optional value $g_p = 4$ Mbit/sec/trk is $\xi = 4$ usable. Thus if we decide against the half-speed playback we should not implement the quadruple fanout either.

Speedup Factor and Channel Sample Rate vs. Playback Sample Rate. A limit on f_r inversely proportional to α is imposed by the limit $f_p \leq 16$ Msmp/sec/chnl; combining this with the provision against undersampling, we have a limit $\alpha \leq 8$ MHz/ B for observing bandwidth B . The correlator specifications call for delay tracking to support $\alpha = 2$ at $B = 8$ MHz and $\alpha = 4$ at $B = 4$ MHz, which is unnecessarily generous. We should reduce the correlator specification to reflect the maximum we can actually record (or perhaps simply drop it). Note, however, that this bandwidth effect does not apply to the required *phase*-tracking accuracy.

IMPLEMENTATION OPTIONS

Half-Speed Playback and Rapid Correlator Dumps. The VLBA correlator will read out its accumulated results at a maximum rate of 2 Hz (in *playback* time) for the entire complement of cross-correlation lags. This dump rate was designed to accommodate (barely) the residual fringe rates occurring in the largest known configurations of compact structures; it coincides conveniently with the largest data rate supportable in the fringe-processor and archive subsystems.

A playback speedup will induce a corresponding decrease of the effective dump rate in *observe* time, and a consequent shrinkage of the Array's field of view. To preclude this shrinkage in observations where an extended field of view is essential, I have defined special widefield modes which reduce the speedup factor (at least) to unity and allow correlation at the maximum effective dump rate. As Table 1 shows, it is always possible to achieve a unit speedup factor (using only the officially-specified full-speed playback) by sufficient oversampling.

However, it is obviously a risky approach to specify a future system on the basis of current astronomical data, and I think we must foresee cases requiring an expanded field of view, and thus effective dump rates exceeding 2 Hz. There are two basic possibilities: implement a fractional "speedup" with a reduced playback speed (an increased record speed has not been considered seriously); or require the correlator to support higher dump rates from some fractional configuration without violating an aggregate dump-rate limit, which would then require multi-pass processing. These possibilities are currently being considered as options which may be implemented in the recording and correlator subsystems.

For technical reasons the reduced playback speed would be limited to the half-speed option considered throughout this memorandum. I have shown several other areas where it could be used to advantage were it available, but clearly the only justification for this option is to raise the effective dump rate. Since this scheme is limited to a doubling of the field of view, it is probably unsatisfactory as a sole expedient — and we should avoid implementing both options! Multi-pass processing is considerably more flexible, but of

course not without its cost to Array operations. The 2-pass case takes no longer than the corresponding half-speed playback, but beyond perhaps 8 passes I doubt that the Array will want to support extremes of this scheme.

It may be useful to consider the several axes along which the correlator could be partitioned for multi-pass processing. The most convenient partitioning is by channels, which probably requires a minimum of sorting of the correlated data. However, high-resolution spectroscopy cannot apply this approach because only one channel is present, and in the low-resolution modes defined earlier only 4 passes are possible. Partitioning by stations into sub-arrays minimizes the wear on tapes and heads, since each tape is only played back once, but invites subtle closure problems (which we hope careful design will exclude). The simplest method in the correlator, partitioning in time by accepting only a fraction of the high-dump-rate correlations, is also the most inconvenient in the post-correlation sorting.

Oversampling Support and Sensitivity. The VLBA recording system specifications require quite generally that narrowband observations ($B < 1$ MHz) be oversampled. Spectroscopic observations in these bands will pose the Array's most severe sensitivity limits; since the oversampling "comes for free" in terms of Array resources, these cases also represent the primary motivation for thorough support of oversampled data in the correlator.

One of the effects of oversampling (discussed above in Section III) is an enhancement of sensitivity when the excess samples are correlated. In the VLBA correlator, an appropriate scaling of the sample shift clock will result in correlation of all, some, or none of these samples. Under the official specifications, however, speeding up the shift clock beyond the Nyquist rate reduces the lag step between successive correlations, and for a fixed number of lags the spectral resolution is degraded proportionally. In this situation, only those few cases where coarse resolution is sufficient could benefit from the forced (or elective) oversampling.

A correlator option under consideration would permit more thorough support of oversampling by providing intermediate lag steps between adjacent correlations to retain full resolution when correlating data oversampled by factors of 2 and perhaps 4. A quantitative consideration of the value of this option should refer to the third column of Table 3 for the relevant enhancement factors. As is well known, more than half the sensitivity improvement occurs at an oversampling factor of 2; I would suggest that oversampling factors of both 2 and 4 yield useful gains and are worth supporting at full spectral resolution unless the marginal cost in correlator hardware is excessive.