Axes of discovery: The time variable Universe

James Cordes∗†
Cornell University
E-mail: jmc33@cornell.edu

As Heraclitus might have said, “You don’t observe the same universe twice,” and in modern times we recognize the time domain as an important dimension in the overall phase space of variables that characterizes the observable universe. Examples abound across the electromagnetic spectrum and in non-photonic regimes (neutrinos, gravitational waves, cosmic rays). However, while we can glimpse the richness of time-domain phenomena at radio wavelengths, the radio sky is largely unexplored in any comprehensive sense, especially when compared to the successes of wide-field surveys at high energies. Known radio transients are as short as 0.4 ns with an equivalent brightness temperature of $10^{42}$ K (Hankins & Eilek 2007) related to the coherent nature of pulsar radiation; others with incoherent emission extend to hour and longer time scales with thermal brightness temperatures. Some time domain properties are intrinsic to sources while others are imposed by multipath propagation through intervening plasma. This paper discusses both known and speculative aspects of the radio transient sky, with an emphasis on discoveries that can be made with new, appropriately designed instrumentation and telescopes. A generalized survey figure of merit is presented that takes into account the rate and duration of transient celestial events. The key for expanding discovery space is a wide field of view (FoV) combined with adequate sensitivity and high-resolution sampling in time and frequency. I discuss implementation of time-domain studies as an integral part of synoptic survey modes and the potential for cross-wavelength and joint photonic/non-photonic studies. In particular, I make the case for designing and operating the mid-frequency range Square Kilometer Array as a Radio Synoptic Survey Telescope.

From Planets to Dark Energy: the Modern Radio Universe
October 1-5 2007
The University of Manchester, UK

∗Speaker.
†Work supported by the US National Science Foundation
1. Introduction

Time variability is a tool for studying astrophysical objects themselves and for probing intervening media and spacetime. The transient sky is on the astrophysical frontier across the entire electromagnetic spectrum and also in non-photonic regimes, including cosmic rays and gravitational waves. The transient sky at high photon energies has been explored much more than in other bands as evidenced by the great successes of wide field-of-view instruments in finding gamma-ray bursts, accreting sources and bursting pulsars over the last forty years. The next large gamma-ray facility, GLAST (the Gamma-ray Large Area Space Telescope) will survey the entire sky every three hours for at least its first year of operation, allowing a comprehensive inventory of both steady and variable sources.

At optical wavelengths, searches for gravitational microlensing events and for Type Ia supernovae have driven the development of high throughput optical telescopes. Forthcoming instruments building upon these successes include PanSTARRs (Panoramic Survey Telescope and Rapid Response System) and the Large Synoptic Survey Telescope (LSST/LST). This paper discusses the transient radio sky. Examples of transient radio signals are known that range from 0.4 ns and longer in time scale with apparent brightness temperatures from thermal to $10^{42} \text{K}$. However, compared to the high-energy transient sky, we know next to nothing about the overall constituency of the transient radio sky. A highlighted science area for the Square Kilometer Array (SKA) is “Exploration of the Unknown” (Wilkinson et al. 2004), which includes the overall phase space opened up by the SKA and the likely discovery of new classes of objects and phenomena. Another chapter in the SKA science book, “The Dynamic Radio Sky” (Cordes et al. 2004) discusses the anticipated payoff from an SKA design that combines widefield sampling of the sky with high sensitivity and flexibility in analyzing likely or hypothetical event signatures and time scales. Most known radio transients have been found by making radio observations on targets selected from surveys made at other wavelengths or energies, such as radio afterglows from gamma-ray burst (GRB) sources and the turn-on of periodic radio pulsations from a formerly quiescent magnetar, J1810-197 (Camilo et al. 2006). Recent exceptions include the discovery of the RRAT population (McLaughlin et al. 2006) and transient sources in the Galactic-center (GC) direction (Hyman et al. 2006; Bower et al. 2006) found through VLA imaging observations of the GC. It is clear that blind surveys of the transient sky will yield new objects and new classes of objects that will serve as laboratories for the basic physics of extreme states of matter as well as inventoring Nature’s proclivity for astrophysical complexity. Indeed, Nature is said to abhor a vaccuum, and this appears to be true in parameter space as well as physical space.

Given the wide diversity of signal types among known radio transients and recognizing that new classes of transients remain to be discovered, our discussion focuses on how best to sample both slow and fast transients in blind surveys. Such surveys are part of the stated science cases for essentially all planned radio telescopes, including the SKA and pathfinders for the SKA and related projects, including the Allen Telescope Array (ATA, Hat Creek), LOFAR (the low-frequency array, 1This particular name is not informative as to the wavelength and would be better changed by adding an “O” for optical to the acronym.}
Nature can be profligate in producing radio photons and thus make sources detectable. It does so in many cases via coherent radiation processes that give $N^2$ rather than $N$ scaling from a collection of $N$ particles. In most sources, radio emission is not the dominant energy channel for dumping free energy but it is an extremely significant information channel. Examples include pulsars and radio pulses from ultra high-energy cosmic rays.

### 3. Transient Phase Space

We know little about the transient radio sky. Giant pulses from radio pulsars are prototypes for fast transients along with solar bursts and flare stars, while sources such as microquasars and gamma-ray burst (GRB) afterglows exemplify longer-duration transients. We may use these sources as initial guides for specifying blind-survey parameters. However, simple observational phase space arguments suggest that *instantaneous coverage of a large fraction of the sky with appropriate sampling of the frequency-time plane will yield a rich variety of transient sources, including new classes of objects.* Figure 1 shows the phase space of pulse width $W$ against flux density. Lines of constant brightness temperature are calculated assuming that $W$ is the light-travel time across the source,

$$T_b = \frac{S}{2k} \left( \frac{D}{\nu W} \right)^2 = 10^{20.5} \text{K} \, S_{\text{mJy}} \left( \frac{D_{\text{kpc}}}{\nu_{\text{GHz}} W_{\text{ms}}} \right)^2$$  \hspace{1cm} (3.1)

where $S_{\text{mJy}}$ is the peak flux density (mJy) at frequency $\nu$ (GHz) and $D$ is the distance (kpc). For some sources $W$ can be much smaller than the light travel time owing to relativistic compression and of course other sources can vary much more slowly.

It is useful to classify the wide range of known time scales in terms of how they are best sampled empirically. *Slow transients* are defined as those with time scales longer than the time it takes to image the relevant region of the sky (e.g. the Galactic plane, the Galactic center, or the entire sky), either in a single pointing or as a mosaic or raster scan. Detection of such objects can be accomplished simply through repeated mapping of the sky and thus does not require special capabilities beyond those needed for imaging applications. Depending on survey speed, transients of days or more may be considered slow. GRBs are currently detected at $\gtrsim 100 \mu$Jy levels using the VLA at frequencies of 5 and 8 GHz (Frail et al. 2006). The full SKA could detect GRB afterglows up to 100 times fainter levels.

*Fast transients*, conversely, are those that would be missed in the time it takes to scan the sky, leading to incompleteness of the survey. Sub-second transients are linked to coherent radiation and, in many cases, to compact sources in extreme matter states by making a simple light-travel
Axes of discovery: The time variable Universe

James Cordes

Figure 1: Time-luminosity phase space for known radio transients, a log-log plot of the product of peak flux $S_{pk}$ in Jy and the square of the distance $D$ in kpc vs. the product of frequency $\nu$ in GHz and pulse width $W$ in s. The “uncertainty” limit on the left indicates that $\nu W \gtrsim 1$ as follows from the uncertainty principle. Lines of constant brightness temperature $T = SD^2/2k(\nu W)^2$ are shown, where $k$ is Boltzmann’s constant. Points are shown for the ‘nano-giant’ pulses detected from the Crab (Hankins & Eilek 2007), giant pulses detected from the Crab pulsar and a few millisecond pulsars, and single pulses from other pulsars. Points are shown for Jovian and solar bursts, flares from stars, brown dwarfs (e.g. TVLM 513-46546; Hallinan et al. 2007), OH masers, and AGNs. The regions labeled ‘coherent’ and ‘incoherent’ are separated by the canonical $10^{12}$K limit from the inverse Compton effect that is relevant to incoherent synchrotron sources. Arrows pointing to the right for the GRB and intra-day variable (IDV) points indicate that interstellar scintillation (ISS) implies smaller brightness temperatures than if characteristic variation times are used to estimate the brightness temperature. The growing number of recent discoveries of transients illustrates the fact that empty regions of the $\nu W - S_{pk}D^2$ plane may be populated with sources not yet discovered. Recent discoveries include the “rotating radio transients” (RRATs; McLaughlin et al. 2006), the Galactic center transient source, GCRT J1745-3009 (Hyman et al. 2006), the bright, bursting magnetar XTE J1810-197 (Camilo et al. 2006), and recently discovered transients (labelled “RT”) associated with distant galaxies (Bower et al. 2007). The rightward-directed triangles used to mark the two RT sources indicate that the transient durations are only known to be longer than $\sim 20$ min.
Possible transients include maximal giant-pulse emission from pulsars, prompt radio emission from GRBs, bursts from evaporating black holes, and radar signals used to track potentially impacting asteroids and comets. Dashed lines indicate the detection threshold for the full SKA for sources at distance of 10 kpc and 3 Gpc. Dotted lines correspond to a 10% SKA, comparable to the Arecibo telescope. At a given $\nu W$, a source must have luminosity above the line to be detectable. The curves assume optimal detection (matched filtering).

Fast transients require the same observing modes and post-processing as for pulsars. The Crab pulsar is the most extreme known case in terms of showing temporal structure down to $\sim 0.4$ ns scales (Hankins & Eilek 2007) and giant pulses that exceed 100 times the entire flux density of the Crab Nebula. Pulsars and giant-pulse emission may represent prototypes for coherent radiation from other high-energy objects in which collimated particle flows can drive the necessary plasma instabilities. Examples include prompt radio burst emission from GRB-type sources, perhaps even from gamma-ray quiet objects; flare stars, jovian-burst like radiation from planets, AGNs, and merging NS at cosmological distances (Hansen & Lyutikov 2001), some of which we designate in Figure 2.

Photon reprocessing may limit the radio brightness of GRBs, so significant non-detections will constrain the source conditions (Macquart 2007).
Using GRBs as a guide, it may be noted that the rate of GRB detection with gamma-ray instruments has relied more on instantaneous wide-field sky coverage than on sensitivity. The same statement holds for the detection of prompt optical emission at $m_v = 9$ in at least one case (Akerlof et al. 1999; Bloch et al. 2000). For this reason, we suggest that use of subarray modes to increase the instantaneous sky coverage — at less than full sensitivity — will be productive in surveying the transient radio sky.

The signal-to-noise ratio for a pulse after matched filtering (including dedispersion for fast transients) is

$$\frac{S}{N} = \frac{S|p_{\text{ol}}BW|^{1/2}}{S_{\text{sys}}}, \quad (3.2)$$

where $S$ is the peak flux density and $W$ is the pulse width (FWHM). Requiring $S/N > m = 10$ and using nominal parameters for Arecibo and the SKA, we get a minimum detectable flux density

$$S_{\text{min,SP}} = \left(\frac{m}{10}\right) \left(\frac{1 \text{ ms}}{W}\right)^{1/2} \left(\frac{1 \text{ GHz}}{B}\right)^{1/2} \times \begin{cases} 21 \text{ mJy} & \text{AO} \\ 0.98 \text{ mJy} & \text{SKA} \end{cases}$$

$$\left(\frac{g_\theta}{g_\theta f_c}\right) \quad (3.3)$$

corresponding to a minimum brightness temperature,

$$T_{\text{min,SP}} = \left(\frac{m}{10}\right) \left(\frac{1 \text{ ms}}{W}\right)^{5/2} \left(\frac{1 \text{ GHz}}{B}\right)^{1/2} \left(\frac{D_{\text{kpc}}}{V_{\text{GHz}}}\right)^2 \times \begin{cases} 10^{21.8} \text{ K} & \text{AO} \\ 10^{20.5} \text{ K} & \text{SKA} \end{cases}$$

$$\left(\frac{g_\theta}{g_\theta f_c}\right) \quad (3.4)$$

A burst of amplitude $S_{pk}$ from a known source at distance $D = D_{\text{kpc}} \text{ kpc}$ can be detected to a maximum distance

$$D_{\text{max}} = D \left(\frac{S_{pk}}{S_{\text{min,SP}}}\right)^{1/2}$$

$$= D_{\text{kpc}} \left(\frac{S_{pk}}{1 \text{ Jy}}\right)^{1/2} \left(\frac{10 \text{ m}}{m}\right)^{1/2} \left(\frac{W}{1 \text{ ms}}\right)^{1/4} \left(\frac{B}{1 \text{ GHz}}\right)^{1/4} \times \begin{cases} 6.9 \text{ kpc} g_\theta^{1/2} & \text{AO} \\ 32 \text{ kpc} (g_\theta f_c)^{1/2} & \text{SKA}\end{cases}$$

$$\left(\frac{g_\theta}{g_\theta f_c}\right) \quad (3.5)$$

The SKA can see about 1.5 orders of magnitude fainter than Arecibo. However, the much greater advantage of the SKA over Arecibo will consist of the sky coverage and greater resilience against radio-frequency interference (RFI). With such coverage we can expect the SKA to yield new discoveries over much of the phase space depicted in Figure 2.

### 4. How Bright Can Fast Transients Be?

Some of the material presented here is from SKA Memo 97, "The Square Kilometer Array as a
Radio Synoptic Survey Telescope: Widefield Surveys for Transients, Pulsars and ETI,” that can be found on the SKA web site\(^2\).

Nature can be prolific in generating bright radio bursts because it doesn’t take much energy to do so and also because coherent emission is easily made at radio wavelengths, so that emission from \( N \) particles goes as \( N^2 \) instead of \( N \).

**Giant Pulses:** We already know that giant pulses from certain pulsars, like the Crab, are emitted frequently enough to be detectable at plausible rates out to 1.5 Mpc with Arecibo. This number results from scaling 0.43 GHz pulses of amplitude \( S_{pk} \sim 150 \text{ kJy} \) that are pulse-broadened to \( \sim 100 \mu \text{s} \) duration. Such pulses occur from the Crab at a rate \( \sim 1 \text{ hr}^{-1} \). They correspond to a pseudo-luminosity \( S_{pk}D^2 \sim 10^{5.8} \text{ Jy kpc}^2 \). Surely there are more luminous, giant-emitting pulsars that are detectable even further.

**Extrapolations:** The case can be made that there are other burst sources that can tap larger sources of free energy than are available in pulsars like the Crab. These alternatives may include

- Pulsars born near the break-up limit (\( \sim 1 \text{ ms} \)) with canonical or magnetar-like magnetic fields (\( 10^{12} \) to \( 10^{15} \text{ Gauss} \)). The rotational energy (using a moment of inertia of \( 10^{45} \text{ gm cm}^2 \))

\[
\frac{1}{2}I\Omega^2 = 2\pi^2 IP^{-2} \sim 10^{51.3} \text{ erg} I_{45}P_{ms}^{-2}, \tag{4.1}
\]

which is comparable to the non-neutrino energy released in a supernova. This energy will be released in a short amount of time as the pulsar rapidly spins down but along the way can drive giant-pulse emission much larger than seen from the Crab, e.g. by a factor of one million, e.g \( S_{pk}D^2 \sim 10^{12} \text{ Jy kpc}^2 \).

- Neutron stars may reactivate their magnetospheres when they merge, with orbital motion substituting for spin in the generation of voltage drops. This kind of system is a probable source for “short” gamma-ray bursts. The energy involved is similar to or exceeds that for extreme giant pulse-emitting objects. The GRB peak luminosity is fiducially \( L_{\gamma} = 10^{51}L_{\gamma,51} \text{ erg s}^{-1} \).

We assume that the true radio luminosity is some multiple \( \varepsilon_r \) of this: \( L_r = \varepsilon_r L_{\gamma} \). Remarkably, over many kinds of astrophysical objects (stars, pulsars, AGNs), we find \( \varepsilon_r \), ranging from about \( 10^{-8} \) (Crab pulsar) to \( 10^{-3} \) (blazars). Here we use a fiducial value \( \varepsilon_r = 10^{-5}\varepsilon_{\gamma,5} \) and a radio emission bandwidth \( \Delta v_r = 1 \text{ GHz} \Delta v_{r,\text{GHz}} \). We then calculate the pseudo-luminosity \( L \) in units of Jy kpc\(^2\) by calculating the peak radio flux from the GRB assuming (only fiducially) isotropic emission. This gives a stupendously large pseudo-luminosity,

\[
L = 10^{15.9} \text{ Jy kpc}^2 \left( \frac{\varepsilon_r \varepsilon_{\gamma,5} L_{\gamma,51}}{\Delta v_{r,\text{GHz}}} \right), \tag{4.2}
\]

corresponding to a peak flux density

\[
S_{pk} = 10^{2.9} \text{ Jy} \left( \frac{\varepsilon_r \varepsilon_{\gamma,5} L_{\gamma,51}}{\Delta v_{r,\text{GHz}}} \right) \left( \frac{3 \text{ Gpc}}{d_L} \right)^2. \tag{4.3}
\]

Beaming may influence the radio luminosity estimate as it does the \( \gamma \)-ray luminosity.

\(^2\)http://www.skatelescope.org
New Sources: Lorimer et al. (2007) detected a 30 Jy pulse at 1.4 GHz that has DM$\sim 375$ pc cm$^{-3}$ from a direction toward which the general interstellar media in the Galaxy and the Small Magellanic Cloud cannot account for the electrons. The authors argue for a 500 Mpc distance based on a 4% baryonic, ionized intergalactic medium. This corresponds to $S_{pk}D^2 \sim 10^{13}$ Jy kpc$^2$, much larger than any Galactic source now known but comparable to the maximal giant-pulse emission discussed above while being somewhat less than our estimate for prompt GRB emission (with admittedly much latitude on scaling parameters).

Scaling from the observed properties, the maximum detectable distance for the event is

$$D_{\text{max}} = D \left(\frac{S_{pk}}{S_{\text{min}}}\right)^{1/2},$$

where

$$S_{\text{min}} = \frac{mS_{\text{sys}}}{(N_{\text{pol}}BW)^{1/2}},$$

and $m = \text{minimum S/N needed for detection}$, $S_{\text{sys}} = T_{\text{sys}}/G$ is the SEFD, $N_{\text{pol}}$ is the number of polarization channels used (1 or 2), $B$ is the bandwidth, and $W$ is the event duration and thus the integration time (for optimal detection). For steady sources, $W$ would be replaced with the dwell time $\tau$ used for a given pointing. For the Parkes survey that found the event, $S_{\text{min}} \sim 0.3$ Jy while for Arecibo it is about 0.03 Jy, in both cases using $\tau = 5$ ms and a detection threshold $m = 10$. This means that $D_{\text{max}}/D \sim 10$ for Parkes and $\sim 32$ for Arecibo.

A perplexing aspect of the Lorimer et al. pulse is that it is a lone event, in two senses. No other pulse was seen from the same pointing direction in many follow-up observations. Nor were weaker — and presumably much more frequent — pulses seen from any sky position, which are expected if the pulse arose from one member of a homogeneously distributed population. The event is strong enough so that a large number of weaker events from more distant sources should have been seen. That none were (pending further reanalysis of Parkes data, PALFA data, etc.) suggests that $D$ may be much smaller than 500 Mpc. If for example, the source is Galactic, then $D \lesssim 10$ kpc and $D_{\text{max}}$ may then be small enough that we should not expect homogeneity to hold. In this case, the absence of weaker events may not be surprising.

5. The Radio Synoptic Survey Telescope

Surveys of time-variable sources need to take into account the event rates and durations as well as sensitivity requirements. Transients cover a very wide dynamic range in time scale and we have distinguished between slow and fast transients. Slow transients are those for which the sky may be sampled by raster scanning, covering a total solid angle $\Omega_t$ in a time $T_t$ and then repeating the scan. The dwell time per sky position is $\tau = (\Omega_t/\Omega_s)T_s$, where $\Omega_t$ is the instantaneous FoV. Figure 3 shows a schematic view. Slow transients are also those for which pulse broadening from multipath propagation is unimportant. Fast transients are those that cannot be well sampled through raster-scan imaging surveys, unless they are very frequent and there is tolerance for a low
Axes of discovery: The time variable Universe

James Cordes

survey completeness. Transients shorter than about one day qualify as fast if an all-sky survey is contemplated. Rare, fast transients are better sampled through staring observations of large solid angles.\(^3\)

5.1 Figure of Merit for Raster-scan Surveys of Steady Sources

For blind searching, the rate of sky coverage \(\Omega\) (deg\(^2\) s\(^{-1}\)) needs to be maximized while also achieving the desired search depth, which we characterize as the maximum detection distance \(D_{\text{max}}\). Survey yield is an obvious metric and it involves the product of source number density \(n_s\), search volume \(V_{\text{max}} = \frac{1}{3} \Omega_s D_{\text{max}}^3\). If the survey covers total solid angle \(\Omega_s\) in a time \(T_s\), the resulting search volume yields a combination of parameters that is the same as obtained by calculating survey speed, \(SS = \Omega_s / \tau\), but taken to a different power. The resulting figure of merit “FoMSS” (which can be read as a FoM for either steady sources or for survey speed) is

\[
\text{FoMSS} = B \left( \frac{N_{\text{FoV}} \Omega_{\text{FoV}}}{N_{\text{sa}}} \right) \left( \frac{f_c A_e}{m T_{\text{sys}}} \right)^2,
\]

where \(N_{\text{FoV}}\) is the number of fields of view (or pixels) for each antenna, \(\Omega_{\text{FoV}}\) is the solid angle for each FoV, \(N_{\text{sa}}\) is the number of subarrays into which the array is divided (assumed equal in size and pointed in non-overlapping directions), \(f_c\) is the fraction of the total effective area \(A_e\) usable in the survey, \(m\) is the threshold S/N in the survey, and \(T_{\text{sys}}\) is the system temperature. This expression is consistent with the simple form often used for survey speed, \(B \Omega_{\text{FoV}} (A_e / T_{\text{sys}})^2\), but makes manifest other variables relevant to surveys.

5.2 Figure of Merit for Pulsars (“Steady Transients”)

Many pulsars, though time dependent, have relatively steady pulse amplitudes. A survey metric for pulsars therefore bears some resemblance to that for steady sources, but the detectability is a function of period \(P\) and dispersion measure \(DM\).

The survey volume is \(V_s = \frac{1}{3} \Omega_s D_{\text{max}}^3\) with \(D_{\text{max}} = (L_p / S_{\text{min}})^{1/2}\) as before, but now the minimum detectable flux density is written as

\[
S_{\text{min}}(P, DM) = \frac{S_{\text{min}_1}}{h_2(P, DM)},
\]

\[
S_{\text{min}_1} = \frac{2 m k T_{\text{sys}} N_{\text{sa}}}{A_e \sqrt{2 B \tau}},
\]

\(^3\)In terms of number of detected events, it is equivalent to cover a large solid angle through raster scanning and small duty cycle per sky position, or to cover a smaller solid angle with continuous time coverage. If the goal is to characterize the angular distribution of the transient population, then raster scanning is appropriate. However, if the event rate of particular sources is desired or if the source population is restricted to a small solid angle (e.g. the Galactic center region), then staring observations are needed.
Figure 3: (Top:) Schematic view of a raster scan where each small square indicates the instantaneously sampled solid angle. The star symbol indicates the occurrence of a transient of duration $W$ at the indicated sky position. (Bottom:) Timeline for the sampling of the event that begins at time $t_{e1}$ during a part of the raster scan when the sky position of the transient is not in the instantaneous solid angle being sampled; the next sky position, that of the transient source, is sampled while the transient is still occurring (shaded rectangle); the transient persists while the scan samples the next sky position.

where the subscript “1” on $S_{\text{min1}}$ denotes the threshold for a single harmonic. The harmonic sum $h_\Sigma$ is given by

$$h_\Sigma(P, DM) = N_h^{-1/2} \sum_{j=1}^{N_h} |\tilde{f}_j|,$$

(5.4)

where $f_j$ is the Fourier amplitude for a time series dedispersed using dispersion measure DM; the harmonic number $j$ corresponds to harmonics $j/P$ of the period $P$; $f_j$ is normalized to the zero-frequency ($j = 0$) value and $N_h$ is the number of harmonics that maximizes the sum; $h_\Sigma$ is equal to unity if the signal is an undistorted sinusoid but can be much larger than unity when the sum is optimized for $N_h \gg 1$. For a gaussian-shaped pulse, this number is $N_h \approx P/2W$ for a period $P$ and pulse width (FWHM) $W$. If the pulse is heavily broadened by instrumental effects, scattering or orbital motion, $h_\Sigma \ll 1$ and the survey becomes insensitive.

Using these definitions, the figure of merit for a pulsar survey is

$$\text{FoMPSR} = \text{FoMSS} \times h_\Sigma^2(P, DM)$$
\[ = B \left( \frac{N_{\text{FoV}} \Omega_{\text{FoV}}}{N_{\text{sa}}} \right) \left( \frac{f_c A_e}{m T_{\text{sys}}} \right)^2 h_\Sigma^2 (P, DM). \]  

(5.5)

Comments:

1. \( h_\Sigma (P, DM) \) is dimensionless and is simply a multiplier of the steady-source FoM.
2. Propagation effects (dispersion and scattering) reduce \( h_\Sigma \) for high-DM objects because pulses are smeared out, particularly at low frequencies; \( h_\Sigma (P, DM) \) is therefore strongly frequency dependent.
3. Application of Eq. (5.5) is best done by averaging over direction, e.g. Galactic coordinates \( \ell, b \), for a particular survey.
4. As presented, \( h_\Sigma \) and FoMPSR are explicit functions of \( P \) and DM. One could alternatively define the independent variables as being \( P \) and distance, \( D \). The corresponding DM will vary greatly for different directions, in this case. In the end it probably makes little difference as to whether DM or \( D \) is used once an average over direction is done.
5. The pulsar FoM, like the other FoMs, implicitly is based on the assumption that the population of sources is homogeneously distributed. Careful application of FoMPSR must recognize the spatial distribution of pulsars in the Galaxy and that highly sensitive surveys will reach the boundaries of the population.

5.3 Figure of Merit for Raster-scan Surveys of Transient Sources

FoMSS applies to surveys of sources that are homogeneously distributed within their spatial domain, that are standard candles, and that are time steady. It is not a good metric for transient sources. A more general metric that also applies to transients is

\[
\text{FoMTS} = \text{FoMSS} \times K(\eta W, \tau/W)
\]

(5.6)

\[
K(a, x) = (1 + x^2)^{-1/2} \left[ 1 - e^{-a(1+x^2)^{1/2}} \right]^{4/3},
\]

(5.7)

where \( a \equiv \eta W \) (a measure of the overlap of events in time) is the product of event rate per source \( \eta \) and event duration \( W \) and \( x \equiv \tau/W = \Omega_i/\Omega W \), with \( \Omega_i = N_{\text{sa}} N_{\text{FoV}} \Omega_{\text{FoV}} \). The quantity \( \Omega = \Omega_i/T_s \) is the mean rate at which solid angle is surveyed. The factors included in \( K(a, x) \) account for the integration time being determined by the transient duration \( (W) \) rather than the raster-scan dwell time \( \tau \) when \( W \ll \tau \) and the probability that a source is pointed at when the event occurs, assuming Poisson statistics. The function \( K(a, x) \) is shown in Figure 4 and is discussed further in SKA Memo 97, Appendix A. The \( K(a, x) \) factor can drastically reduce the effective survey speed from what it is for steady sources.

5.4 Completeness Coefficient of a Transient Survey

Transient sources by definition are “on” and hence detectable for only a fraction of the time, which may be small. We therefore need to consider how a survey samples a population of transient sources in both the spatial and temporal domains. A completeness coefficient \( C_s \) can be defined as the ratio
of number of objects that are detected to the number of objects that could be detected during the scan time $T_s$. For a homogeneous population of sources distributed over solid angle $\Omega_{\text{pop}}$ and out to a distance $D_{\text{pop}}$, this takes on the form

$$C_s = \min \left[ \frac{\Omega_s}{\Omega_{\text{pop}}}, 1 \right] \min \left[ \left( \frac{D_{\text{max}}}{D_{\text{pop}}} \right)^3, 1 \right] \frac{P_t(\eta, W, \tau)}{P_t(\eta, W, T_s)},$$  \hspace{1cm} (5.8)

where $P_t$ is the temporal capture factor (the probability that at least one event occurs when the telescope is pointed at a bursting source), given by (for Poisson statistics)

$$P_t(\eta, W, \tau) = 1 - e^{-\eta (W^2 + \tau^2)^{1/2}}.$$  \hspace{1cm} (5.9)

There are three regimes for $C_s$ that depend on the event rate $\eta$ and on whether the events are fast or slow: $\eta W \ll \eta \tau \ll 1$, $\eta \tau \ll \eta W \ll 1$ and $\eta (W^2 + \tau^2)^{1/2} \ll 1$. If there are very bright, but rare, prompt radio bursts associated with GRBs, we expect $P_t \approx \eta \tau$, and $D_{\text{max}}$ to equal the population
distance, in which case \( C_s \approx 10^{-4} \) for a 10m antenna outfitted with a single-pixel feed. To detect such bursts requires much larger field of view.

5.5 Trading Field of View and Sensitivity

The sensitivity and FoV requirements for transients can be discussed in terms of the event rate for particular classes of sources. Figure 5 shows minimum detectable flux density plotted against instantaneous solid angle. For single-reflector, single-pixel telescopes, we have a simple relationship,

\[
S_{\nu,\text{min}} = \frac{mS_{\text{sys}}}{\sqrt{2B\tau}} = \frac{8kT_{\text{sys}}}{\eta_A\pi\sqrt{2BT}} \left( \frac{m\Omega_{i,1}}{\lambda^2} \right),
\]

(5.10)

where the unity subscript implies a single aperture and we have assumed that \( \Omega_{i,1} = (\lambda / D)^2 \) (i.e. unity coefficient). The integration time is \( \tau \), which may or may not be related to the dwell time on a given source position. The factor \( m \), as before, is the threshold in units of the rms noise level for a bandwidth \( B \) and integration time \( T \). The solid line in Figure 5 shows this relationship. For arrays with \( N_A \) antennas each having a multi-pixel receiver with \( N_{\text{FoV}} \) pixels, we have

\[
\Omega_i = N_{\text{FoV}}N_{\text{sa}}\Omega_{i,1}
\]

(5.11)

\[
S_{\nu,\text{min}} = \frac{1}{N_A N_{\text{FoV}}} \frac{8kT_{\text{sys}}}{\eta_A\pi\sqrt{2BT}} \left( \frac{m\Omega_i}{\lambda^2} \right).
\]

(5.12)

If \( N_{\text{sa}} \) subarrays are used, the FoV increases at the expense of higher threshold. We show points in the figure for the full SKA outfitted with a single-pixel system (SKA/SP) and with a 32-pixel system (SKA/PAF). We also show a point for an SKA built around an aperture array, which provides ~ 1 sr FoV. The slanted dashed line shows the relationship for subarrays, the ultimate configuration comprising \( N_A \) subarrays of one antenna each. The figure delineates the various permutations of strong/weak and rare/common transients. Vertical dashed lines indicate the solid angle required to capture sources of indicated event rates in a reasonable dwell time (see caption). In the figure we keep the integration time, \( T \), fixed, as appropriate for short-duration transients with \( W \ll \tau \), i.e. \( W \) shorter than the dwell time per position, in which case \( T = W \). To have the temporal capture probability \( P_t \sim 1 \) we require, for low-rate, short-duration transients \( W \ll \tau \),

\[
\Omega_i \gtrsim \frac{1}{(\eta / \Omega_s)T_s}.
\]

(5.13)

In the figure we have considered event rates \( \eta / \Omega_s \) of 0.1 s\(^{-1}\) deg\(^{-2}\) with a one-hour scan time and 1 day\(^{-1}\) deg\(^{-2}\) and 1 day\(^{-1}\) hemisphere\(^{-1}\) with scan times of one day for the three vertical, dashed lines from left to right. The rightmost line, for example, is relevant to the case of prompt radio pulses from GRBs, should they exist.

From the figure we can make a number of conclusions. First, if there are hyper-strong events (such as coherent radio pulses from GRBs), a wide FoV system is necessary for a reasonable detection rate, but modest collecting area may suffice for detection. However, rare weak events require high sensitivity as well as wide FoV. Field of view expansion of single reflectors allows probing of the upper right-hand corner as does an aperture array. But it is also true that wide FoV may be achieved
Figure 5: Plot of minimum detectable flux density ($S_{\nu, \text{min}}$) vs. instantaneous solid angle $\Omega_i$. The solid diagonal line refers to single-reflector telescopes with single pixel feeds and all having — for simplicity in presentation — the same center frequency (1 GHz), operating bandwidth (0.3 GHz), system temperature (25 K) aperture efficiency (60%), and integration time (1 s). The dashed extension of the single reflector line denotes reflector diameters < 6$\lambda$. The “Full SKA” line indicates the sensitivity for 4400 12-m antennas. “FoV expansion” implies an increase in solid-angle coverage through use of dishes + phased-array feeds (PAFs) or through an aperture array (AA). The subarray line indicates the tradeoff between sensitivity and instantaneous FoV for single-pixel systems. Vertical dashed lines indicate the solid angle needed to detect events at the indicated rates assuming a one day total exposure time. The ellipse in the bottom-right corner indicates the fiducial amplitude of hypothetical prompt radio GRB emission. The location of the GRB ellipse implies that dipole antennas may yield sufficient solid angle coverage but will be too insensitive to detect prompt GRB emission for nominal values of the burst parameters.
at moderate sensitivity through the use of subarrays. Since the low-event-rate radio sky is largely unexplored, we should entertain the possibility of using the subarray approach as well as develop FoV expansion approaches.

### 6. AS vs. SA: Aperture Synthesis vs. Subarrays

It is generally more productive to use an array in aperture synthesis (AS) mode than in extreme subarray (SA) mode. Consider the case where a total time $T$ is used to observe a region of size $\Omega_s$. In AS mode, we fast-sample and mosaic a region of size $\Omega_s = N_p \Omega_1$ with integration time $\tau = T/N_p$ per pointing. Alternatively in an extreme subarray (“fly’s eye”) mode with one antenna per subarray, we would stare continuously at the same region of the sky for the same total time. Therefore the number of pointings $N_p$ for AS equals the number of antennas, $N_A$. The number of events detected for transients that are fast in both AS and SA mode ($W \ll T/N_p \ll T$) is

$$\frac{N_e^{(SA)}}{N_e^{(AS)}} = N_A^{-3/2}$$

(6.1)

while for the case where the events are fast for SA mode ($W \ll T$) but slow for AS mode ($W > T/N_p$) we have

$$\frac{N_e^{(SA)}}{N_e^{(AS)}} = N_A^{-3/2} \left( \frac{W}{\tau} \right)^{3/4}.$$  

(6.2)

#### 6.1 The Synoptic Cycle: An Example

The observing cycle for a given survey consists of a raster scan of a region of sky $\Omega_s$ in a time $T_s$, yielding a dwell time per sky position $\tau = T_s \Omega_i/\Omega_s$. This cycle can repeated many times, allowing detection of a growing number of transient sources, both periodic ones like pulsars and magnetars, and one-shot objects like GRBs. At the same time, signal-to-noise ratio is built up for steady sources, such as HI in galaxies, continuum sources for AGN surveys and Faraday rotation measurements, etc. Because multiple surveys are potentially doable, and each is demanding on total telescope time, all effort should be made to do the surveys simultaneously (commensally).

To accomplish multiple surveys, a hierarchical approach to scanning rates (or cadences) is probably needed. A fast rate is appropriate for the extragalactic sky in order to sample GRB afterglows while also building S/N on HI galaxies. A slower rate is needed for the Galactic disk to provide adequate time series durations for pulsar surveys. Staring observations of the Galactic center source will allow deep pulsar and transient surveys from the star cluster. Finally, guest-investigator experimental or one-time observations need to be accommodated along with target-of-opportunity observations that will arise.

An example scenario includes both fast and slow scanning observations and staring observations. This example sums to about 10 days per cycle, which would be repeated as needed:
1. **Fast scan of the extragalactic sky:** large-scale galaxy HI survey, Faraday rotation survey, AGN survey, and transients
   
   (a) “Full sky” survey (80% of $4\pi$) using a 1 deg$^2$ FoV single pixel system
   
   (b) $T_s = 5$ days to cover one scan of the sky
   
   (c) $\tau \approx 10$ s dwell time per sky position
   
   (d) $S_{\text{min}} \approx (g_\theta f_c)^{-1} 15$ $\mu$Jy at $10\sigma$ where $f_c$ is the fraction of a full SKA available and $g_\theta \leq 1$ is the gain relative to the on-axis gain
   
   (e) Field-of-view expansion through multiple feed clusters or phased-array feeds will increase the sensitivity for fixed $T_s$; aperture arrays would also provide an increase. A thirty-beam phased-array feed, for example, would yield 300 s dwell time
   
   (f) Subarrays will reduce the sensitivity but can cover more instantaneous solid angle and reduce $T_s$.

2. **Slower scan or staring observations on deep extragalactic fields** e.g. 1 day

3. **Slower scan for the Galactic plane:** pulsars, masers, transients, etc. e.g. 1 day scan of the inner Galaxy (180 deg in longitude) in a ±1 deg swath in Galactic latitude, yielding 240 s per pointing
   
   (a) Minimum contiguous dwell time needed for pulsar surveys that use Fast Folding Algorithms or Fourier transforms of a contiguous time series combined with harmonic summing (100 to 1000 s typical); single-pulse searches do not place strong requirements on contiguous blocks.
   
   (b) Pulsar timing: frequent re-observations are needed for long-term monitoring; a 10-day cadence is acceptable.

4. **Staring observations:** e.g. 12 hr on the Galactic center

5. **Break out for targeted observations by individual investigators:** 10% of the time?

6. **Break out for targets of opportunity:** e.g. GRB triggers, blazar observations, etc. 5% of the time?

7. **Calibration allowance**

Additional comments on this scenario are as follows:

1. HI detection of galaxies at $z \sim 1$ requires many hours of integration time, which would build up slowly unless there is field-of-view expansion.

2. Pulsar surveys can accumulate S/N through incoherent summing of power spectra from non-contiguous data segments, with due allowance for acceleration of the source or observer. Coherent sums across multiple days are probably too demanding computationally, but with requirements that depend on the cadence.

3. Diffractive and refractive interstellar scintillations will modulate compact sources to varying degrees and with a wide range of correlation times. To optimize detection, multiple passes on the same sky position should be uncorrelated with respect to DISS and RISS (Cordes & Lazio 1991).
6.2 Analyzing the Frequency-time Plane

All observations, whether of steady sources or the fastest transients, require processing of signals in both time and frequency. The common thread across all source types includes calibration and RFI excision. RFI, in particular, must be sampled with high resolution in the $\nu - t$ plane in order to identify its various types (narrow/broad band, impulsive, swept-frequency, etc.) and excise it.

Signal detection invariably makes use of matched filtering or some approximation to it. A matched filter (MF) by definition maximizes the signal-to-noise ratio of the test statistic used to define detection. MFs are used for identifying sources in images, spectral lines in spectra, pulses in time series, and more complex events in spaces of higher dimensionality. Familiar examples include identification of dispersed pulses in the $\nu - t$ plane and spectral lines with variable Doppler shifts from acceleration of the source or the observer. The MF is usually parametric; when values of parameters are not known, the MF becomes a family of filters whose performance is optimized over a grid of values.

Figure 6 shows examples of signal types in the frequency time plane. One basic point is that, phenomenologically and topologically, a dispersed or otherwise frequency-swept signal can appear as a drifting spectral line and vice versa. An important difference of course is that the drift rates of frequency-swept pulses are typically much larger than the variable Doppler drift rates of spectral lines.

**Pulse and Spectral Line Detection:** For pulse detection in a time series, $W_p$ is the characteristic width of the pulse and $W_n$ is the characteristic time scale of the noise, comparable to the sample interval in digitized data that is Nyquist sampled. By inspection, the S/N in the time series, $S_{pk}/\sigma_n$, increases by a factor $(W_p/W_n)^{1/2}$ in the MF test statistic, in accord with the expected $\sqrt{N}$ law. For a spectral line, $W_p$ and $W_n$ are characteristic frequency scales.

**Dispersed Pulses:** When a pulse propagates dispersively, its arrival time varies with frequency, $\Delta t = 4.15 \text{ ms DM} \nu^{-2}$ for dispersion measure DM in standard units (pc cm$^{-3}$) and $\nu$ in GHz. In the frequency-time plane, the signal shape $I(t, \nu)$ has a template that depends on the dispersion measure, DM, and on the temporal pulse width. “Post-detection” dedispersion sums over $\nu$ taking into account dispersive delays. The remaining step of matched filtering involves smoothing of the resulting time series with a filter that matches the pulse shape, $p(t)$.

A more exact method is “coherent” dedispersion, where the signal model is for the electric field rather than intensity and thus involves phase as well as amplitude. Coherent dedispersion involves correction of the phase wrap of the signal that corresponds to the frequency-dependent time delays in the post-detection method. As with the post-detection method, the coherent dedispersion MF is a one-parameter filter (DM); signal detection then involves smoothing with $p(t)$, as before.

In survey applications, DM is not known a priori, so it becomes a parameter in a family of MFs. In applications on objects with DM known to arbitrary precision, coherent dedispersion may be viewed as exact while post-detection dedispersion is only an approximation of varying convergence to the exact case.
Figure 6: Schematic views of six types of signals as seen in the $\nu - t$ plane. From left to right, top to bottom we have: (a) a single cell with resolutions $\Delta \nu$ and $\Delta t$ in an observational unit that spans total bandwidth $B$ and total time $T$; (b) a drifting event with instantaneous widths $W_\nu$ and $W_t$ and total spans $B_e$ and $T_e$. For a uniform drift rate $\dot{\nu}$ we have $W_\nu/W_t = B_e/T_e = \dot{\nu}$. Events of course may have curved trajectories in the $\nu - t$ plane. (c) a time-steady spectral line; (d) a broadband pulse; (e) a drifting spectral line with steady amplitude; (f) a dispersed pulse, including curvature of the path according to the cold-plasma dispersion law. Drifting structures as in (b), (e) and (f) may occur from processes other than those mentioned (variable Doppler shift, dispersive propagation), including refraction effects, stellar and Jovian bursts, and gravitational shifts.
**Pulse broadening:** Multipath propagation from scattering and refraction in the ISM causes intrinsic pulse shapes \( p_i(t) \) to be convolved with an asymmetric function \( p_{\text{MP}}(t) \) to produce \( p(t) = p_i(t) \ast p_{\text{MP}}(t) \). As above, the matched filter for detection is the measured shape \( p(t) \).

**Periodic Dispersed Pulses:** Periodicity introduces the period \( P \) into the overall MF. For a train of pulses with fixed \( P \) and equal amplitudes, the MF is \( \sum_j p(t - jP) \). The filter output over one cycle corresponds to the “folded” pulse shape often computed in pulsar applications after smoothing by \( p(t) \) to obtain maximal S/N. Search applications require a family of such rail filters with parameter \( P \). An approximate method — widely used — consists of Fourier transformation of the time series followed by summing of harmonics for trial periods \( P \) and trial numbers of harmonics. The resulting S/N is sub-optimal because harmonic summing typically does not include phase information but can approach the MF result to within a factor \( 2^{-1/4} \).

**Orbiting pulsars:** Orbital motion causes received pulses to be unequally spaced. For orbital periods much longer than the data span of interest, a single parameter is needed — the acceleration — in the MF, which is essentially a rail filter with unequally spaced pulses. To search for objects with short orbital periods, a five Keplerian parameters need to be included in the MF. Approximate methods include standard Fourier analysis on short data sets and summing of resultant spectra or analyzing orbital sidebands in the Fourier transform of the entire time series (ref).

**Precessing pulsars:** Slowly precessing pulsars will mimic orbiting pulsars, to some extent. Objects that have complex, triaxial precession, may tumble and will defy a MF approach to an entire time series. Detection may need to rely on single-pulse detection.

**ETI carrier signals:** One type of proposed ETI signal is a simple carrier that drifts in frequency from accelerated motion by a drift rate \( \dot{\nu} \). The MF takes into account the line width and drift rate. If the signal is transient with duration \( T_e \), that too must be included in the MF.

**Faraday rotation:** Rotation of the plane of polarization by an angle \( \chi \propto \nu^{-2}\text{RM} \), where RM is the rotation measure. Detection of the linearly polarized flux density is accomplished by derotating the quantity \( L = Q + iU \) in the complex plane by a phase \( \phi = 2\chi \) with RM as an unknown parameter.

**Acknowledgments**


**References**


