



Artificial Broadcasts as Galactic Populations. I. A Point Process Formalism for Extraterrestrial Intelligences and Their Broadcasts

Brian C. Lacki^{1,2}

¹ Breakthrough Listen, Department of Astronomy, University of California Berkeley, Berkeley, CA 94720, USA; astrobrianlacki@gmail.com

² Breakthrough Listen, University of Oxford, Oxford OX1 3RH, UK

Received 2022 June 30; revised 2023 November 30; accepted 2023 December 1; published 2024 May 7

Abstract

Artificial broadcasts from extraterrestrial intelligences are a hypothetical class of celestial phenomena. Unlike known astrophysical objects, the societies that generate them may be able to replicate on galactic scales through interstellar travel. Different galaxies could thus have drastically different populations, with variations in abundance of many orders of magnitude. I present a probabilistic formalism to treat this shared history, in which societies and their broadcasts are described by distributions over basic properties like lifespan and energy released. The framework contains a hierarchy of objects related by a tree structure. Discrete societies, the sources of broadcasts, are organized into potentially interstellar *metasocieties*. The population of each type of object is represented by a random point process in an abstract parameter hyperspace, a *haystack*. When a selection like an observation draws a sample, the point process is thinned. Given assumptions of interchangeability and independence, observables are modeled with compound Poisson random variables. I present an example of how selection bias can favor sampling longer-lived objects. I rederive the Drake equation for societies in the limit of no expansion. When interstellar replication is present, however, the mean number of detected broadcasts can depend quadratically on stellar mass, suggesting a search strategy favoring large galaxies.

Unified Astronomy Thesaurus concepts: Search for extraterrestrial intelligence (2127); Technosignatures (2128); Astronomical techniques (1684); Poisson distribution (1898); Spatial point processes (1915)

1. Introduction

The search for extraterrestrial intelligence (SETI; Tarter 2001; Worden et al. 2017) rests on the premise that the technosignatures of extraterrestrial intelligences (ETIs) are astrophysical phenomena, arising in and shaped by a cosmic environment and detectable over cosmic distances. ETIs are not ruptures in the laws of physics but consequences of them, forming a cosmic population, however rare they actually are. The properties of this population are constrained by physical limitations on technology and unknown internal factors.

Negative SETI results, although leaving plenty of room for ETIs, are at least strong enough to eliminate many possible classes of broadcasts as the main contributor to the Milky Way's luminosity. Narrowband radio beacons with isotropic luminosities of around 10^{12} – 10^{26} W cannot possibly make up a significant fraction of its current radio emission (Enriquez et al. 2017; Price et al. 2020; Tremblay & Tingay 2020; Włodarczyk-Sroka et al. 2020; Gajjar et al. 2021; see Paper II). Optical SETI constraints also are sensitive to energy outputs far smaller than the Galaxy's stellar luminosity (Horowitz & Sagan 1993; Tellis & Marcy 2017; Maire et al. 2019). Very little work has been done in SETI at higher energies like X-rays (Corbet 1997; Hippke & Forgan 2017), or alternate messengers like neutrinos (Subotowicz 1979; Learned et al. 1994; Hippke 2018). But our ability to interpret the cosmos as natural implicitly suggests that artificial broadcasts are not the main emission process in other wave bands (admittedly, not a certainty as per Čirković 2018a).

If we were looking for conventional sources around stars, we would conclude that the brightest examples of the source class

are rare in all galaxies as they should more or less just trace stellar mass, with additional biases from things like metallicity and age. Only exotic environments like globular clusters and galactic nuclei might have much higher abundances than seen in the field, as for close binary star systems, stellar mergers, and their products (Bailyn 1995; Pooley et al. 2003; Muno et al. 2005; for ETIs, see Di Stefano & Ray 2016).

Unlike all known astrophysical phenomena, however, ETIs potentially *replicate*.³ Interstellar travel can allow intelligent beings who evolved on a single world to spread to a large fraction of a galaxy in well less than 1 billion yr if such propagation is practical (Hart 1975; Tipler 1980; Jones 1981; Zackrisson et al. 2015; Carroll-Nellenback et al. 2019). Widespread replication presents great opportunities for SETI, as it vastly increases the technosignature profile of a galaxy (Kuiper & Morris 1977; Wright et al. 2014). We have no evidence that this has happened in our Galaxy, a statement known as the Fermi paradox (Brin 1983; Wright et al. 2014; Webb 2015; Čirković 2018b; Forgan 2019; Lingam & Loeb 2021).

However, this does not rule out it happening in other galaxies. If ETIs with interstellar travel are rare but not impossible, we could end up with a situation where most galaxies (perhaps including our own) have none, but a few have millions of inhabited worlds. Galaxies that look alike to us could have divergent technosignature populations due to unobservable historical factors (perhaps set by the first ETIs as in Scheffer 1994 and Hair 2011). The menagerie of societies in galaxies populated by starfarers did not all arise independently, but have a shared origin in perhaps a single world, even if the resulting population is extremely diverse. This shared history can shape the entire population of descendant ETIs, and so the group might be treated as a whole—neither a simple collection of noninteracting societies, nor a generic galaxy solely described by



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

³ Life might also replicate on interstellar scales via panspermia (e.g., Napier 2004).

Universe

Galaxies

Metasocieties

Societies

Broadcasts

Emission

Measurements

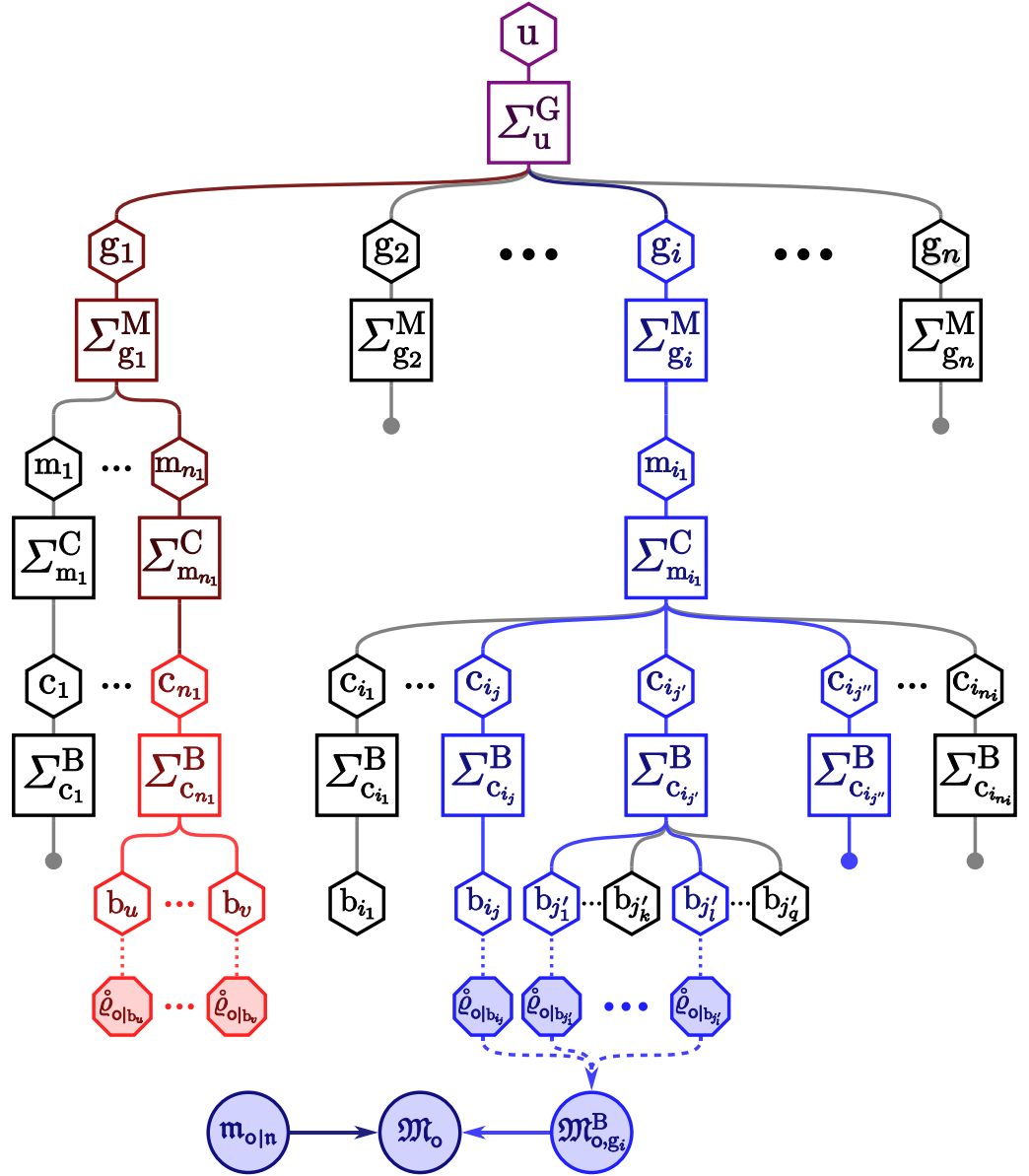


Figure 1. Overview sketch illustrating the tree structure of objects in this series. A selection draws samples of galaxies, metasocieties, societies, and broadcasts. The galaxy g_1 is in the classic SETI scenario with isolated worlds, each developing its own metasociety, while g_i has a single galaxy-spanning metasociety, and g_2 and g_n are uninhabited (additional objects are represented by the ellipses). Here, an observation (\mathcal{O}) of a galaxy (g_i) defines a more restrictive selection (\mathcal{O} , g_i) that picks a subsample of objects, highlighted in bright blue. The windowed emission $\mathcal{M}_{\mathcal{O},g_i}^B$ from the broadcast sample $\Sigma_{\mathcal{O},g_i}^B$ is mixed with background ($m_{\mathcal{O}|n}$) into an observed quantity, $\mathcal{M}_{\mathcal{O}}$. The selection also picks the ancestors of g_i (dark blue). Some societies and broadcasts in g_i are excluded because they lie outside the \mathcal{O} window (in a different field or at a different time, for example). The objects and emission sampled by another selection, (∞ , c_{n1}), are highlighted in red (dark red for ancestors of c_{n1}).

astrophysical properties. I call these intermediate-level groupings *metasocieties*.⁴ Therefore, the underlying parameters of the technosignature distribution in each galaxy are themselves the result of a stochastic process, relating to the properties of the underlying metasociety. As a result, constraints on ETIs in one

galaxy could mean little for the presence or absence of ETIs in another.⁵ SETI has set the first limits on technosignatures in other galaxies (e.g., Horowitz & Sagan 1993; Shostak et al. 1996; Annis 1999; Griffith et al. 2015; Gray & Mooley 2017; Garrett & Siemion 2023), but these limits still lag far behind those in the Milky Way.

Since ETIs could be biased toward specific galaxies where interstellar travel has flourished, this suggests a treatment of galactic ETI populations and their technosignatures is in order. We can use probability theory to calculate the expected emission

⁴ Compare analogous notions of *metapopulations* and *metacommunities* in ecology, both describing situations in which organisms are clustered in many small-scale habitats that influence each other through dispersal—members of a single species for metapopulations (Hanski 1998), and sets of multiple interacting species for metacommunities (Leibold et al. 2004). Metapopulation ecology bears a direct analogy with several of the dynamics of interstellar migration and metasocieties—for example, the survival of a metapopulation depends on quality and number of habitat patches (see Carroll-Nellenback et al. 2019). The analogy between interstellar migration and metapopulations is explicitly noted in Lingam & Loeb (2021), who argue it implies that not all worlds will be inhabited at any given time.

⁵ Neglecting the possibility of intergalactic travel, which could lead to cosmological bubbles within which ETIs may synchronize (Kardashev 1985; Armstrong & Sandberg 2013; Olson 2016). Even then, different domains of the Universe could have varied technosignatures.

Table 1
Summary of Fonts Used

Font	Case	Examples	Use	Frame
Blackboard bold	Upper	$\mathbb{R}, \mathbb{M}, \mathbb{V}$	Certain sets and operations	...
Hebrew	...	$\aleph(K), \beth(B)$	Object type operations	...
Roman	Any	W, χ, δ	Mathematical functions and probability distributions	...
	Lower	j, m, b	Indices for realized objects	...
San-serif	Upper	J, G, MW	Labels for types of objects, random objects, and certain named objects	...
	Lower	μ, o, x, t	Labels for windows	...
Italics	Upper	B_o, Π_x, S_y^x	Variables defining a selection window	Source
	Any	$X, t, f(x), k_B, H_0$	Generic variables and functions; physical and cosmological constants	...
	Lower	$\kappa_{x j}, w_u, \vartheta_m, \hat{e}_b$	Quantities describing intrinsic properties of individual objects	Source
	Upper	$\Psi_k^j, M_g^*, \Sigma_X^M, \hat{L}_{\nu,\nu'}^B$	Quantities describing samples of objects	Source
Fraktur	Lower	d_b, z	Distances, redshifts, dilutions, and transmittances	...
	Lower	$m_{x b}, h_{o B}$	Fluence, flux, or observables for a single object or background	Observer
	Upper	$\mathfrak{M}_x, a_{\mu n}$	Fluence, flux, or observables from a sample, or in total	Observer
Calligraphic	Upper	$\mathcal{A}, \mathcal{N}_A$	Quantities describing instrument	Observer

Note. Some characters, mainly the Greek letters, have no counterparts in the fonts used. For these, the standard italic or upright character is used. A $^{\oplus}$ superscript attached to a source-frame variable changes it to an observer-frame variable.

from all the ETIs covered by an observation. Developing this population approach is the focus of this paper and its sequels.

1.1. The Problem of Notation

The tables in this paper list the important variables. There are many kinds of objects with their own populations and derived constraints (Figure 1). Each of these levels has its own associated parameters, which leads to a great proliferation of variables. Additionally, many of them already have often conflicting symbols in the literature. This can easily lead to a highly overloaded namespace: for example, many different types of quantity may lay claim to N —number of stars, number of societies, number of broadcasts, number of polarizations emitted, number of pointings, number of photons counted, and so on. Subscripts and superscripts help, but risk confusion when two completely different types of variables share the same symbol (e.g., L for luminosity and longevity in the Drake equation), and lead to clutter when combined with the selection notation developed later. I use different fonts to help distinguish between different types of variables, as listed in Table 1. The chosen notation does not always match their usage in other fields.

1.2. Outline of Paper

I begin the paper with a review of point process theory (Section 2). The construction of selections and distributions is detailed in Section 3. The nature of randomness in ETIs and broadcasts is discussed in Section 4. A discussion of the modeling of the universe, galaxies, and their stellar populations (Section 5) is followed by the treatment of the interplay of ETI metasocieties and societies (Section 6). Next is consideration of the properties of individual broadcasts (Section 7), and the results are assembled into calculations of broadcast populations and their total emission (Section 8). Sections 9 and 10 present the box and chord models of broadcasts for calculating observables. A brief discussion of noisy measurements comes in Section 11. Section 12 demonstrates the formalism by considering how a galaxy’s stellar mass affects the number of broadcasts in the face of sporadic interstellar travel. After the conclusion (Section 13), appendices present additional details and derivations.

2. Background on Point Processes

Point processes have a rich theoretical background. The concepts throughout this section are informed by Kingman (1993), Daley & Vere-Jones (2003, 2008), Baddeley (2007), Chiu et al. (2013), Haenggi (2013), and Last & Penrose (2017). The reader may consult them for further details. Certain technicalities are given in footnotes for interested readers.

2.1. Random Variables

Probability is defined over a sample space of all outcomes, each of which is a distinguishable configuration of all variables. Because outcomes are so particular, it is practical to group them into events defined by a shared quality. An event’s probability $P(E)$ is summed (or integrated) over all outcomes in the event E (Wasserman 2004). Outcomes can be summarized by a random variable, which assumes one value per outcome called a variate. Formally speaking, a random variable is defined as a function mapping each outcome onto a set, conventionally a real number (Klenke 2020). Let us say we are interested in the positions of stars in the Milky Way. Every single possible configuration of stars is a distinct outcome. Those outcomes with two stars within 1 parsec of the Earth form an event; the number of stars within 1 parsec of the Earth is a random variable, which depends on what outcome is realized.

Distributions describe the relative probability of different variates, among them the cumulative distribution function (CDF) $F[X](x) \equiv P(X \leq x)$ and probability density function (PDF) $\psi[X](x) \equiv dF[X](x)/dx$ for real-valued variables (Wasserman 2004)⁶ (refer to Table 2 for the statistical notation used in this series).

Random variables are very general, and can have completely arbitrary values. Among the simplest is the indicator variable: $\mathbb{I}[E]$ is 1 if event E happens and 0 if it does not. Random variables can have a deterministic value regardless of outcome, like always being 0, with a degenerate distribution. Thus, random variables per se do not imply a lack of rhyme or reason.

⁶ The PDF is conventionally written as f_X . I use ψ to avoid confusion with other uses like the factors in the Drake equation.

Table 2
Summary of Statistical Notation Used

Notation	Explanation
Variables	
X	Generic random variable
N	Generic nonnegative integer random variable
$\mathbb{I}[E]$	Indicator random variable; is 1 if the event E happens and 0 otherwise
$X Y$	Random variable X , with outcomes limited by the value of random variable Y
$\kappa_{x j}$	Arbitrary singleton random variable describing object j ; quantity filtered by window x
$\kappa_{x y,K}^j$	Arbitrary aggregate random variable, summing $\kappa_{x j}$ from j -type objects selected by (y, K)
$ S $	Cardinality of set S , when S is nonnumeric
Probability Distributions and Operations	
$X \sim Q(\alpha)$	X has Q distribution with parameter α
$P(E)$	Probability of some event E
$\odot[X]$	Any operation on the distribution of the random variable X : can stand for event probability, CDF, PDF, characteristic function, mean, variance, standard deviation, and order statistics including maxima and minima
$\odot[X]_{x,K}$	Ensemble selection-relative distribution operation on X_j for a random object j drawn by (x, K) (Section 3.4)
$\odot[\kappa_{x j}]_{z y,K}$	Mult>window distribution operation on $\kappa_{x j}$ that lets $\odot[\kappa_{x j}]_{y,K}$ range over all the x_i windows in z (Section 3.5).
$P_{x,K}(E_j)$	Ensemble probability of an event related to random object j selected by (x, K)
$F[X](x)$	CDF of X , $P(X \leq x)$
$F[X Y](x y)$	Conditional CDF of X conditionalized on Y , $P(X \leq x Y \leq y)$
$F[X, Y](x, y)$	Joint CDF of X and Y , $P(X \leq x \cap Y \leq y)$. Can be extended to arbitrarily many variables.
$\psi[X](x)$	PDF of X , $dF[X](x)/dx$. Conditional and joint PDFs substitute ψ for F in the CDF notation.
$\phi[X](\bar{x})$	Characteristic function of X
$\kappa_{x j}^{[y,K]}$	Regularization (trimming) of $\kappa_{x j}$ to exclude values unlikely to occur in $\Sigma_{y,K}^j$
Means and Variances	
$\langle X \rangle$	Mean (expectation value) of random variable X ; simple mean for variables describing an object
$\mathbb{V}[X]$	Variance of random variable X ($\langle X^2 \rangle - \langle X \rangle^2$); simple variance for variables describing an object
Order Statistics	
$\mathbb{M}[X]$	Median of random variable X
$X_{(j)}$	Order statistic: j th smallest in the set of random variables X_k ; distinct from unsorted X_j
$X_{(1)}$	Minimum of the set of random variables X_k
$X_{(N)}$	Maximum of the set of N random variables X_k , where N itself is a random variable

Random variables are independent if we learn nothing new about the value of one if we know the value of the other. Often, however, random variables depend on an outside factor, without influencing each other. Conditional independence of X and Y means that, if we know the value of a third random variable Z , knowing what X is tells us nothing more about the value of Y . It neither requires nor implies independence.⁷ Conditionalization is very important in this series because it can isolate the shared influences between random variables. The conditional mean $\langle X|Y \rangle$ and variance $\mathbb{V}[X|Y]$ come up frequently. The law of total expectation states that

$$\langle X \rangle = \langle \langle X|Y \rangle \rangle, \quad (1)$$

and the law of total variance is

$$\mathbb{V}[X] = \mathbb{V}[\langle X|Y \rangle] + \langle \mathbb{V}[X|Y] \rangle \quad (2)$$

(Brillinger 1969; Wasserman 2004; Bas 2019).

2.2. Point Processes

An astrophysical object may be described by a tuple w of quantitative parameters. Each possible tuple is a point in the set of all tuples, the state space W . Even objects that are extended

in an observable space are represented as points, their internal structure following from the point's location—a core method in stochastic geometry. As the common metaphor goes, our observations trawl through that space to find the “needles in the haystack” (Zwicky 1957; Harwit 1981; Tarter 1984; Djorgovski et al. 2013; Wright et al. 2018).⁸

A population is represented as a point process. Basically, a *point process* is some random set Σ of points on a state space W .⁹ A random variable $N(A) = |\Sigma \cap A|$ counts the number of points appearing in each region $A \subset W$.¹⁰

A core property of each point process is the mean number of points¹¹ in each region A . The means $\langle N(A) \rangle$ are described by a

⁸ In this series, these state spaces are those describing physical properties of objects, not the observable parameter spaces of Harwit (1981) and Djorgovski et al. (2013) describing observational capabilities.

⁹ Each bounded region has only a finite number of points (almost surely), which is no issue for realistic populations. A point may be sampled multiple times, requiring some way to encode multiplicity. However, I assume the point processes are *simple*, each point included only once with probability 1 (Daley & Vere-Jones 2003; Baddeley 2007).

¹⁰ Formally, each region must be a Borel set (Daley & Vere-Jones 2003). Any rigorously defined subset of \mathbb{R}^n prone to be encountered in astronomy is Borel: all continuous curves and surfaces, including fractals, whether unbounded or closed; their interiors and exteriors; Cantor dusts; all open sets; all closed sets; and countable unions of these, among others.

¹¹ Both $N(A)$ and $\langle N(A) \rangle$ are measures. One consequence is that they are additive when taking the (countable) union of disjoint subsets.

⁷ Specifically, it means $F[X, Y|Z](x, y|z) = F[X|Z](x|z) \cdot F[Y|Z](y|z)$ (Wasserman 2004; Klenke 2020).

density distribution Ψ known as the intensity:

$$\langle N(A) \rangle = \int_A \Psi(\mathbf{w}) d\mathbf{w}. \quad (3)$$

The intensity does not have to be a smooth function; it can include Dirac distributions to indicate exact locations where some points are found. Points can also fall along lines or surfaces in the space—just like how the properties of galaxies and stars can be idealized to fall along lines or surfaces in graphs describing their properties, like the main sequence or the Tully–Fisher relation.

Often, we wish to calculate the sum of random variables associated with each point \mathbf{w} in the point process Σ . Campbell’s formula is a very general result that relates it to the distribution

$$\left\langle \sum_{\mathbf{w} \in \Sigma \cap A} f(\mathbf{w}) \right\rangle = \int_A f(\mathbf{w}) \Psi(\mathbf{w}) d\mathbf{w}. \quad (4)$$

There are many ways to manipulate point processes. The *superposition* of several given point processes includes all the points from each of them in a new point process, with the $\Psi(\mathbf{w})$ summing together.¹² *Thinning* a point process keeps some of the points based on their locations in the space. Independent thinnings keep or remove a point solely based on its own location and not on any other points. We can also use a point’s position to *map* it to another location in the space.

Finally, we can *mark* the points in a point process, attaching a random variable with additional information. If the marks are drawn from set \mathbb{K} , a marked point process is equivalent to a point process on $W \times \mathbb{K}$. Furthermore, if the mark κ of a single point is drawn from the distribution $\psi[\kappa|\mathbf{w}]$, independent of the number of other points and their locations, then the intensity on this new space is $\Psi(\mathbf{w}) \times \psi[\kappa|\mathbf{w}]$. That is why the haystack concept works—although objects are located in physical spacetime, their parameters serve as *marks*, which can then be interpreted as additional dimensions in an abstract space.¹³

2.3. Poisson Point Processes

Poisson point processes are special point processes with many useful properties (Kingman 1993).¹⁴ They have seen use in modeling populations of ETIs (Glade et al. 2012; Kipping 2021), fast radio bursts (Lawrence et al. 2017), and galaxies (Neyman & Scott 1952; Martínez & Saar 2002). In a Poisson point process, each of the $N(A)$ has a Poisson distribution of

$$P(N(A) = n) = e^{-\langle N(A) \rangle} \frac{\langle N(A) \rangle^n}{n!}. \quad (5)$$

Additionally, the number of points in nonoverlapping regions are independent random variables. A Poisson point process is

¹² As long as the collection of superposed point processes is countable, there can even be infinitely many of them.

¹³ The reverse operation—treating one or more of the dimensions as marks for a point process on a lower-dimensional space—is only allowed if the resulting projected process does not have infinite $\langle N(A) \rangle$ for any bounded A (Baddeley 2007), but happens whenever we simplify distributions by marginalizing over a parameter.

¹⁴ It is common to use “Poisson process” to refer to the specific case counting the number of hits along a nonnegative real line. A “compound Poisson process” then measures the accumulated sum of jumps occurring at each hit (e.g., Ross 1996; Wasserman 2004; Embrechts et al. 2013).

completely specified by its intensity.¹⁵ The Poissonian character of the point processes is unaffected by superposition, independent thinning, mapping, and displacing points by a random offset. The resulting intensities are what one naively expects—added together for superposition, reduced by a position-dependent fraction in independent thinning, preserved in the image of a subset for mapping, and blurred for displacement—subject to certain technical conditions (Kingman 1993). A marked Poisson process with points on W and marks from \mathbb{K} is equivalent to a new Poisson point process on $W \times \mathbb{K}$, as long as the marks are mutually independent. Other point processes converge to Poissonian when repeated randomizing operations are applied to them, most notably displacing their points independently and randomly (Daley & Vere-Jones 2008).

These results formalize common sense notions in a very general way. Often in astronomy, we model a field of objects with a Poisson process, like counting stars in a patch of sky or transients observed by a radio telescope. These results show that we can add information, like the colors and magnitudes of stars, and still use the Poisson distribution. Poisson statistics also apply when we change the parameterization (e.g., frequency to wavelength), combine or filter out subclasses of objects (like early- and late-type stars), or have scatter (as from time delays or errors). There only has to be a determined intensity (Section 2.5).

2.4. Compound Poisson Variables and Point Processes

A compound Poisson random variable has the form

$$K = \sum_{j=1}^N \kappa_j, \quad (6)$$

where the κ_j are independent and identically distributed (i.i.d.), and also independent of $N \sim \text{Pois}(\langle N \rangle)$ (e.g., Adelson 1966; Barbour & Chryssaphinou 2001; Karlis & Xekalaki 2005; Geringer-Sameth et al. 2015). If $N=0$, then $K=0$. The distribution of K has a compound form as the weighted sum of the distribution for each fixed N :

$$P(K \leq x) = \sum_{n=0}^{\infty} P(N = n | \langle N \rangle) P(K \leq x | N = n). \quad (7)$$

Although this distribution often has no closed form expression, its Fourier transform, the characteristic function, is

$$\phi[K](\tilde{x}) = \langle e^{iK\tilde{x}} \rangle = \exp[\langle N \rangle (\phi[\kappa_j](\tilde{x}) - 1)], \quad (8)$$

where $\phi[\kappa_j]$ is the characteristic function of the κ_j (Kemp 1967; Daley & Vere-Jones 2003). An inverse Fourier transform then gives a numerical probability distribution (Geringer-Sameth et al. 2015). Equation (8) appears in $P(D)$ analyses, which implicitly use the compound Poisson distribution (Scheuer 1957). Both the mean and variance have simple

¹⁵ Additionally, a Poisson point process has $\langle N(\{\mathbf{w}\}) \rangle = 0$ for any single point \mathbf{w} (there are no *atoms* in the intensity; Kingman 1993). It is fine for lower-dimensional curves and surfaces to have $\langle N(A) \rangle > 0$, however (Daley & Vere-Jones 2003).

expressions:

$$\begin{aligned}\langle K \rangle &= \langle N \rangle \langle \kappa_J \rangle \\ \mathbb{V}[K] &= \langle N \rangle \langle \kappa_J^2 \rangle\end{aligned}\quad (9)$$

(Adelson 1966; Bas 2019).

Whenever we have a Poisson point process Σ , with each point marked by i.i.d. nonnegative real numbers, the compound Poisson point process describes the sum of the marks for different subsets $K(A)$ (Daley & Vere-Jones 2003; Last & Penrose 2017). Each $K(A)$ is a compound Poisson random variable. In fact, even the κ_j themselves can be compound Poisson random variables, representing hierarchical clusters of events, as long as they are i.i.d. and independent of $N(A)$ on each level.

What if the marks are independent but their distribution varies with \mathbf{w} ? Although the sum is no longer compound Poisson, Campbell's formula for the mean (Equation (4)) still applies, and a second Campbell's theorem gives us

$$\mathbb{V}[K(A)] = \int_A \langle \kappa(\mathbf{w})^2 \rangle \Psi(\mathbf{w}) d\mathbf{w} \quad (10)$$

if $\langle K(A) \rangle$ exists (e.g., Kingman 1993). This equation is valid only for Poisson point processes, failing when the points are dependent or $N(A)$ is fixed.

2.5. Other Point Processes

Actual counts may not be Poissonian. The binomial point process is what we get when the total number of points in a Poisson point process is fixed. The positions of the points are mutually independent, although the numbers of points in each $A \in \mathcal{W}$ are dependent.

While the Poisson point process is a mixture of binomial point processes with random N , the Cox point process is a mixture of Poisson point processes where the intensity distribution Ψ itself is random. Only one intensity field is ever realized for the point process. Looking at it, we would only see that realized intensity, and it would look like a Poisson point process, no matter how much we sampled it. Cox point processes are thus *nonergodic*. Only if we had a family of identical Cox point processes could we see the true range of variability in them. Cox point processes have found use in modeling ecological populations, when the configuration of offspring is hypothesized to cluster around parents (e.g., Wiegand & Moloney 2013). The populations of societies and broadcasts in galaxies are regarded as the Cox point process later in this paper as well.

Even more general are point processes where the random points interact. At the deterministic limit are lattices, with points repelling one another. Another example, the Gibbs point process, has a global energy function penalizing some configurations of points (e.g., forbidding those with points too close together), but practical results are difficult to come by. They remain outside of this work's scope.

In this work, objects of a single type usually appear randomly and do not interact (with some exceptions described in Section 6.3). Poisson point processes are natural when we have objects appearing independently according to a well-defined background rate. Several kinds of object populations behave this way (see discussion in Section 4), if their immediate host has specified properties, including broadcasts in a society. But sometimes we wish to consider the population hosted by a randomized collection of hosts. The intensity then is itself

random, and the point process is Cox. The locations of stars in a specific galaxy are a Poisson point process (e.g., Tonry & Schneider 1988), with the brightness profile serving as intensity. However, the star counts of a galaxy in the abstract are not defined because the brightness profile itself varies with galactic mass, type, and so on—it is more accurately described as a function that returns a Poisson point process given the galactic parameters. The star counts of a galaxy cluster are a Cox point process since the gross properties of the cluster do not necessarily tell us about the characteristics of the member galaxies, only their *distribution*. The star counts are a superposition of Poisson processes for all possible configurations of galaxies in the cluster. Broadcasts in galaxies are one example of a Cox point process in this paper because they are clustered into random societies.

3. Defining Selections and Samples

We do not observe the properties of ETI broadcast populations directly, only measure limited samples of the entire population. Measurements are treated as random variables that depend on, but are not determined by, the properties of the sampled broadcasts. This introduces two forms of variance: the sampling variance due to the randomness of which objects are included, and noise variance introduced by microscopic fluctuations in observables for that sample (Figure 2). Describing how the sampling is modeled, and how it affects random variables, is the focus of this section.

3.1. Basic Structure of Populations

3.1.1. The Tree and the Haystacks

A theory of the population of ETI technosignatures encompasses the habitat that hosts the ETIs, the ETIs themselves, and their technosignatures. In this formalism, there are two kinds of structures relating them.

First is a multilevel *tree*, assigning each object to a host. The ancestors of an object are its host, the hosts of the host, and so on, with the parent being the immediate node one level closer to the root. Likewise, each object can host a subpopulation of descendants, with the children being those that are one level deeper.

Five *levels* are considered in this series, each for a different type of object. The root is the *universe* (type U), containing all other objects. On the next level down are *galaxies* (type G). These contain *metasocieties* (type M), a collection of societies with a common origin or influence, which may be localized or extended. Within them, (*communicative*) *societies* of ETIs (type C) may send out *broadcasts* (type B). The objects in each level form a population of objects (Figure 1).

The levels are related by type of operations. The ancestor-type $\aleph(J)$ returns the set of object types that are ancestors of J, with the parent-type operation $\aleph_1(J)$ returning the type one level up. The descendant-type $\daleth(J)$ gives the set of object types that are lower on the tree, regardless of whether any such descendants exist, and the type of operation $\beth(J)$ (or $\beth(j)$) that returns an object's type. Thus, for the metasocietal type M, $\aleph(M) = \{U, G\}$, $\aleph_1(M) = G$, $\beth(M) = M$, and $\daleth(M) = \{C, B\}$.

The other ingredient is a series of state spaces for objects, a collection of haystacks (see Wright et al. 2018). Every type of object J has a haystack W^J (Figure 3). A single object J is described fully by a corresponding tuple \mathbf{w}_J in its haystack. Each tuple is a Cartesian product of parameters describing the object. These can include basic quantities like the lifespan of the object, its location, energetics, and so on.

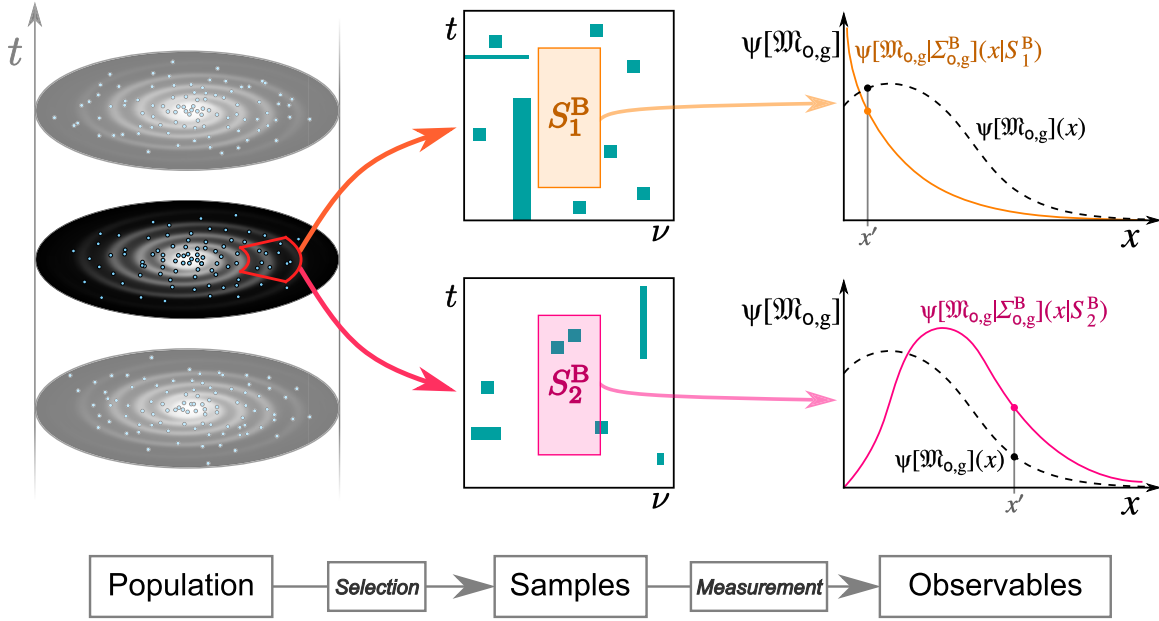


Figure 2. The ETI broadcasts of a galaxy over its history form a population described by distributions (left). Surveys, and observations within the surveys, select subsets of the positions, times, and frequencies (and polarizations, not shown) spanned by this population, drawing samples treated as random sets (middle). In turn, measurement yields observable quantities, which are themselves random variables with probability distributions dependent on the sample drawn by the observation (right). Both the sampling itself and the measurement conditionalized on the observed sample contribute variance to the final quantity.

Included among the intrinsic properties of the object are parameters describing the statistical properties of its internal child population. So, for every object k of type K with $J \in \mathcal{T}(K)$, w_k specifies the *distribution* of the descendants on the corresponding haystack W^J , but generally not which w_J are included in the actual subpopulation.¹⁶

3.1.2. Models

A model describes the populations within some object of interest. It picks some node K in the tree, whether the universe or a single host that we are studying, to serve as the root of a subtree. The model specifies enough parameters that the descendant populations of that node are statistically characterized. The model can be identified with the tuple w_K describing this subtree root object K . Thus, the model is essentially a generalized random variable. The actual population of objects that exists is a realization of the model.

Only one set of model parameters actually describes the model's root. We may know some of these parameters—the stellar mass of a model galaxy, or the cosmological parameters of the universe—but others, particularly those describing ETI subpopulations, are unknown. However, the random nature of the model root is different than the randomness of the descendant populations. The distributions of objects on lower levels have a frequentist character. The subpopulation of J objects within a host K is a single trial; if we had an infinite number of trials, with an ensemble of hosts all with the same w_K , the mean density of J

objects in the haystack really would converge to Ψ_K^J . We use w_K to predict the properties of the descendant subpopulation.

In contrast, there is only one root object, as is clear when it is the universe as a whole, and cannot repeat trials. Yet we are uncertain about the true values of w_K , so we have to consider many possible models, and compare the kinds of populations that result with reality. To the extent there is any distribution for the root object, the probabilities involved are Bayesian, describing our level of belief in each hypothetical model; the goal is inference, not prediction. The results of SETI surveys are phrased as statements about likelihood, the consistency of each model with the data. The classic rate-versus-luminosity plots of SETI (e.g., Enriquez et al. 2017) are an example, describing the consistency with an empirical null result.

3.1.3. Random and Realized Objects

Random variables describe the observables of objects, but sometimes we need a general result—the object itself is a random variable of sorts. These are the random objects, in opposition to realized objects. With a random object, we mean *for anything of this type*, but for a realized object, we mean *this (possibly hypothetical) object*. Samples draw random objects from the haystack, where the object's tuple itself is the random variable.

Each dimension of a haystack is a coordinate variable; realized objects have specific values for each coordinate, while random objects leave them as free parameters.¹⁷ Every quantity associated with a realized object that is not a deterministic

¹⁶ This framework echoes others developed in the past literature across science, without directly following them. In the Neyman–Scott process, clusters of galaxies are represented by random points in real space, and associated with each is a collection of galaxies represented by points randomly distributed by some distribution. The clusters themselves can be clustered this way, to arbitrarily high order (Neyman & Scott 1952, 1958; Martínez & Saar 2002; Haengi 2013). Also related are multiplicative population chains, modeling the growth of population between generations (Moyal 1962a, 1962b; Daley & Vere-Jones 2003). Unlike those formalisms, however, the underlying type of state space changes between generations because each level is a new type of object.

¹⁷ The interpretation of the basic parameters as coordinates and not random variables is implicit in their direct use in the distributions (e.g., $d\langle N_{K,K}^J \rangle / d\tau_J$). The same is true of deterministic combinations of the basic parameters (like broadcast isotropic luminosity from total energy release and lifespan), which can be regarded as using a different coordinate system. However, these intrinsic quantities also could be interpreted as random variables with degenerate distributions, as in Section 3.3.1. The notation for these basic parameters is the same as for the random quantities that mark each notation, however, reflecting this ambiguity. If we wished to be absolutely rigorous, we would make a distinction between the basic parameter as coordinate variable and a random variable wrapper, but I ignore the distinction to avoid further clutter.

Object type of population in haystack

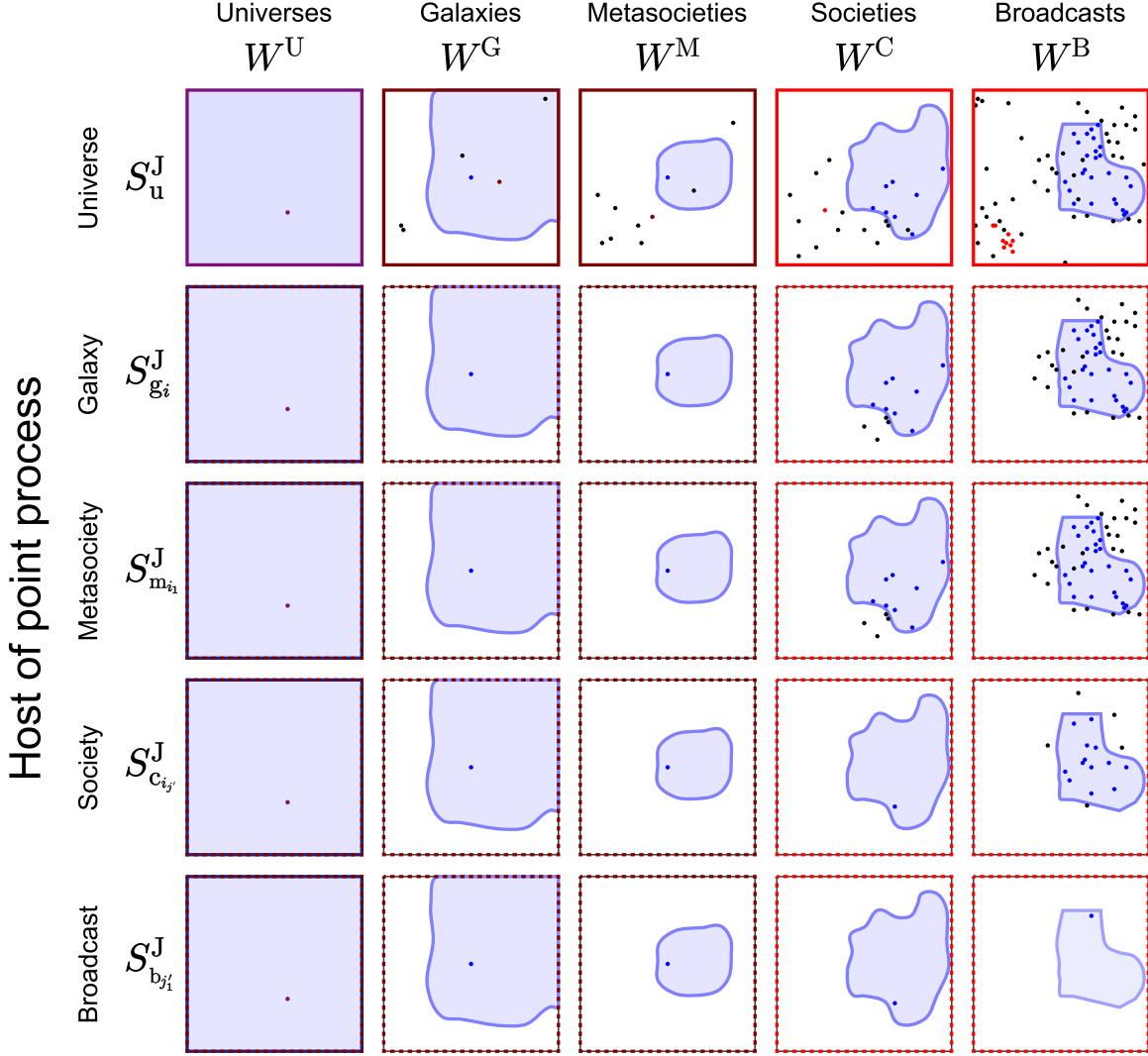


Figure 3. Example illustration of the various haystacks for (left to right) the universe, galaxies, metasocieties, societies, and broadcasts, depicted here as having only two dimensions. Each host, chosen along one branch of Figure 1, has its own point process for each type of haystack. The realized samples illustrated here represent one possible population, each object a single point. Those objects that are selected by (\mathbf{x}, g_i) (blue region) or (∞, c_{i_1}) (red outline) are colored blue and red, respectively. In some cases, an object falling in the \mathbf{x} window is missed because its ancestors on the tree (its hosts) are not selected (black dots in blue regions).

combination of these coordinate parameters is a random variable, with a distribution that is fixed by their values. With random objects, these quantities are not formally random variables, as we cannot specify a single distribution without the object's properties being known. Instead, they actually are *functions* that yield random variables when the tuple describing a realized object is fed into them as input. The number of photons intercepted from random broadcast B during window \mathbf{x} , $q_{o|B}$, is an example of one of these random functions, and should in all strictness be written as $q_{o|B}(\mathbf{w}_B)$. When \mathbf{w}_B is assigned a single value \mathbf{w}_b , then $q_{o|b} \equiv q_{o|B}(\mathbf{w}_b)$ is a bona fide random variable with a single distribution. In practice, the random functions may be written as if they are variables to avoid cluttering notation, but the dependence is still implicitly there.

The notation makes this distinction in the way it indexes random variables. An uppercase index designating a type of object also refers to a random object of that type. When the index is a lowercase letter, it denotes a realized object. Thus, for a random variable X and a J-type object j described by a

parameter tuple \mathbf{w}_j ,

$$X_j = X_j(\mathbf{w}_j). \quad (11)$$

This distinction is very general, applying even to distributions, point processes, and the selection-relative operations—so $\Psi_K^J(\mathbf{w}_J|\mathbf{w}_K)$ is the distribution of J-type objects in a random K-type host, while $\Psi_k^J(\mathbf{w}_J)$ is the distribution in a realized host k .

Most of the results that follow apply to both random and realized objects, and can be interconverted between the two by substituting notation and using the appropriate dependence on the random object's tuple:

$$\begin{aligned} \Psi_K^J(\mathbf{w}_J|\mathbf{w}_K) &\rightarrow \Psi_k^J(\mathbf{w}_J) = \Psi_K^J(\mathbf{w}_J|\mathbf{w}_K = \mathbf{w}_k) \\ \psi[\kappa_{x|J}]_{y,K}(\kappa) &\rightarrow \psi[\kappa_{x|J}]_{y,k}(\kappa) \\ &= \psi[\kappa_{x|J}]_{y,K}(\kappa|\mathbf{w}_J; \mathbf{w}_K = \mathbf{w}_k) \\ \kappa_{x|J}(\mathbf{w}_J) &\rightarrow \kappa_{x|j} = \kappa_{x|J}(\mathbf{w}_J = \mathbf{w}_j) \\ K_{x,K}^J(\mathbf{w}_K) &\rightarrow K_{x,k}^J = K_{x,K}^J(\mathbf{w}_K = \mathbf{w}_k). \end{aligned} \quad (12)$$

3.1.4. Populations as Point Processes

The population of each kind of J object is described by a point process on its haystack W^J (Figure 3). Furthermore, a point process exists for every possible host. The point processes are equivalently described by a random set on the haystack space, Σ^J , or by its resultant collection of $N^J(A)$ (Section 2.2). Only one of the possible sets is actually drawn, the realized sample or sample variate S^J .

Let K be the parent of J-type children. These children form a random set $\Sigma_K^J(w_K)$ on W^J , the J-haystack.¹⁸ So the broadcasts produced by a realized society c are characterized by a point process Σ_c^B on the broadcast haystack W^B , for instance. Of course, the children of an object are constrained by the properties of the parent. Descendant objects cannot be born before their ancestors, for example. As a result, the region of the haystack where children objects may be found is restricted.

Each point process has an intensity distribution of

$$\psi_K^J(w|w_K) = \frac{d\langle N_K^J(w_K) \rangle}{dw_J}(w), \quad (13)$$

where N_K^J is the number of J-type objects descended from the random host K.¹⁹ This equation applies generally, but it serves as a definition for the distribution of child objects in a parent. Because the characteristics of the parent determine the statistical properties of the child subpopulation, the value of the intensity function at each point w_J in the J-haystack is specified by the model and w_K . Note that these mean numbers refer to the population of objects over the entire history of the parent, *not* the typical number of objects that are active at any one time (contrast the N in the Drake equation). Finally, the distribution is an average over an ensemble of all possible realizations of the subpopulations in K.

Of course, we are not just interested in the subpopulation in the parent, but also in the higher-order ancestors. We often wish to know the total number of broadcasts in an entire galaxy, not just one society, for instance. These populations are just the superposition for the point processes of descendant hosts—for an L-type random host with $L \in \aleph(K)$ and $K \in \aleph(J)$,

$$\begin{aligned} \Sigma_L^J(w_L) &= \bigcup_{w_K \in \Sigma_L^K(w_L)} \Sigma_K^J(w_K) \\ N_L^J(A|w_L) &= \sum_{i=1}^{N_L^K(A|w_L)} N_{K_i}^J(A|w_{K_i}) \end{aligned} \quad (14)$$

for each region $A \subset W^J$. The number of K-type subancestors is itself a random variable, and $\Sigma_L^K(w_L)$ is random too. Now, the mean value of $N_L^K(A|w_L)$ is closely related to $\psi_L^K(w_L|w_L) = d\langle N_L^K(w_L) \rangle / dw_L$ through a derivative. The distribution is thus also a weighted sum over the possible K

¹⁸ Or, to be pedantic about it, $\Sigma_K^J(w_K)$ is a random function that yields a random set (point process) based on w_K ; for the realized host k , $\Sigma_k^J \equiv \Sigma_K^J(w_K = w_k)$ is a random set (Section 3.1.3).

¹⁹ The haystack is multidimensional, and the tuples are more or less vectors with many components. Even some of the parameters themselves are multidimensional (e.g., spatial position). The number of dimensions depends on the type of object and model, however; for an n -dimensional variable x , d/dx should be understood to be the n th order derivative, and dx as a hypervolume element.

hosts. The intensity for the total J object population is

$$\psi_L^J(w_J|w_L) = \int_{w_K} \psi_K^J(w_J|w_K) \psi_L^K(w_K|w_L) dw_K. \quad (15)$$

This follows essentially from applying Campbell's formula (Equation (4)) because $\psi_K^J(w_J|w_K)$ is a random variable at each w_K . This way, we can build up the distribution functions from the parent all the way to the universe this way, given a model (see Moyal 1962a). The distribution of objects in their parent-type hosts is fundamental. Given these, we can derive all the other distributions for objects in ancestors, informing us about the populations we expect to be contained within each host.

What if the host object type is not ancestral, though? The J-type population in a J-type object is the self-population, namely, itself: $S^J(w_J) = \Sigma^J(w_J) = \{w_J\}$. It is a Dirac point process, picking the object with certainty

$$\psi_J^J(w|w_J) = \delta(w - w_J). \quad (16)$$

What about the point process of ancestors contained within an object? It helps to distinguish between what ancestors an object *could have* had, and what ancestors an object *actually* has. Only the latter, the realized sample, is used, and it only contains one point for each level. When k is a K-type ancestor of j , $S_j^K = S_k^K = \{w_k\}$.

3.2. Selections

Any realistic observation program cannot observe the entire realized population of objects throughout the universe, much less all possible objects in the haystack. An observation, survey, or other selection instead draws a sample of them based on position, time, frequency, and other constraints, as well as which ancestors they have (Figure 2). A *selection* makes these cuts. Each selection (x, J) is the combination of a *window* x and a *host* object J . Notation for windows is listed in Table 3.

3.2.1. Windows

A window x selects things that pass cuts on variables like location, lifespan, or frequencies covered. They include the reach of a survey or observation, but the cuts can also be more abstract, picking all broadcasts that are active at an arbitrary frequency, for example. A window is thus a kind of selection defined on the haystack of each object type.

Each window acts to thin a point process according to position in the haystack. A window x has two functions. First, it includes a filter on observable quantities, generally a bounded region over some basic variables so that any object that touches it is selected (see Section 3.3.1). Second, a window restricts object selections: formally it includes a collection of probability functions $P_x^J(w_J)$ for each type of object J , giving the probability that the window selects an object with tuple w_J . Basically, P_x^J gives the completeness of the window selection for an object with w_J . All selected objects fall within a subset of the haystack,²⁰

$$W_x^J = \{w_J \in W^J : P_x^J(w_J) > 0\}. \quad (17)$$

Whether an object J is selected by the window depends solely on $P_x^J(w_J)$. It does not otherwise matter whether an object is

²⁰ Windows do not have to be contiguous, but that is assumed in the box and chord models.

Table 3
Notation for Selection Windows

Notation	Explanation
(\mathbf{x}, J)	The selection picking objects and emission using window \mathbf{x} from a population hosted by object J
Useful Windows	
$\mathbf{x}, \mathbf{y}, \mathbf{z}$	Denote generic windows; can be substituted with any other unless otherwise indicated.
μ	Picks objects and emission covered by in one electromagnetic field mode (see Paper II)
\mathbf{o}	Picks objects and emission covered by an observation
θ	Picks objects and emission covered by a single beam or resolution element with a fixed sky position in a series of observations
ν	Picks objects and emission covered by a <i>tuning</i> , a series of observations covering a fixed frequency range (e.g., in a single filter in optical)
π	Picks objects and emission covered by a pointing, a series of observations over a fixed sky field
\mathbf{s}	Picks objects and emission covered by a survey
ρ	Function returning a window that picks all objects and emission coincident with a fixed value of an arbitrary basic variable ρ
\mathbf{t}	Picks all objects and emission active at one instant as viewed from Earth (i.e., along our past light cone); a function of t
\mathbf{h}	Picks all objects and emission in past history, within our past light cone; a function of t
ν	Picks all objects and emission at one source-frame frequency, regardless of time; a function of ν
$\mathbf{t} \cap \nu$	Picks all objects active at one instant as viewed from Earth at one source-frame frequency; a function of t and ν
∞	ALL window: picks all objects and emission, regardless of parameters
Quantities Describing Window Definition	
T_x	Duration of window \mathbf{x}
B_x	Bandwidth covered by \mathbf{x}
Θ_x	Central time of \mathbf{x}
Υ_x	Central frequency of \mathbf{x}
Δ_x	Drift rate of \mathbf{x} window; used for describing dedrifted observations
Π_x	Set of polarizations covered by \mathbf{x}
$ \Pi_x $	Number of independent polarizations covered by \mathbf{x}
Ω_x	Sky field covered by \mathbf{x}
V_x	Spatial volume covered by \mathbf{x}
Window Manipulation	
S_y^x	Set of windows of type \mathbf{x} that are fused into \mathbf{y}
N_y^x	$ S_y^x $, number of component windows of type \mathbf{x} making up \mathbf{y}
$\mathbf{x} \cup \mathbf{y}$	Fused selection, inclusive of both \mathbf{x} and \mathbf{y}
$\mathbf{x} \cap \mathbf{y}$	Joint selection, intersection of \mathbf{x} and \mathbf{y}
$\mathbf{x}; \rho, \mathbf{j}; \rho$	Projection of \mathbf{x} or \mathbf{j} into basic variable ρ ; picks all objects and emission that have a ρ value in the range covered by \mathbf{x} or \mathbf{j}

detectable or distinguishable from a background. For example, the leakage from a cell phone call in a distant galaxy would be selected by a survey window if the survey looks at that galaxy at the right frequency when it happens, unless the window also imposes a luminosity or fluence cut.

A few fundamental variables—position, time, and polarization—define an arena in which things are situated (Table 4). Windows and objects may cover extended regions in this space (Figure 4), and their arrangement is a problem of stochastic geometry. A basic kind of window is one that picks objects along a fixed value of one of these variables. For each variable

Table 4
Notation for the Basic General Variables

Notation	Explanation
t	Time
ν	Frequency
p	Polarization
d	Distance
\mathbf{r}	Position in space
θ	Location on receiver sky
ρ	Arbitrary basic variable

Note. Vectors are shown in bold.

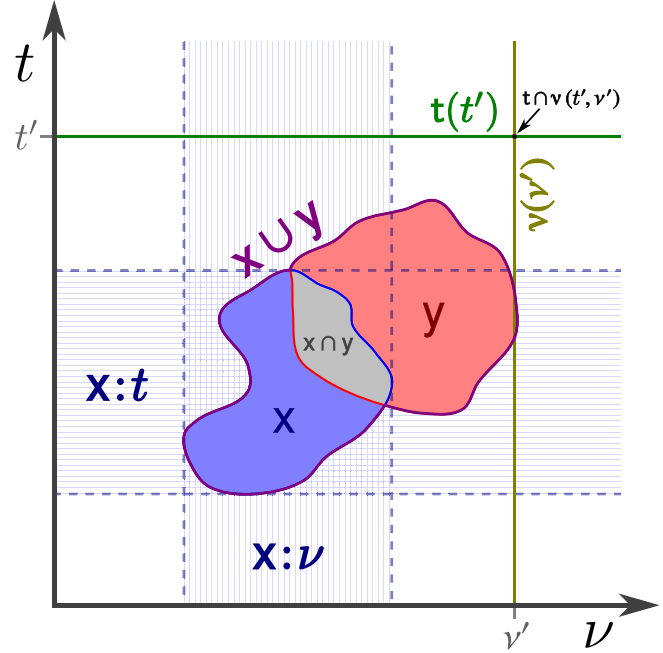


Figure 4. Illustration of how different selection windows could manifest in a spectrogram. window \mathbf{x} (shaded blue) can be projected onto time ($\mathbf{x};t$) and frequency ($\mathbf{x};\nu$). A second window \mathbf{y} (shaded red) can be fused with \mathbf{x} (violet outline), or a joint selection can be made (gray shading). Also shown are windows picking a specific time (green line), frequency (gold line), or both (black dot).

ρ , a function $\rho(\rho)$ returns a window that intersects with a given fixed value of it. The most commonly used one is the time window function, $\mathbf{t}(t)$, selecting every object that exists during a particular moment t in time as viewed from Earth.

In this work, an observation \mathbf{o} is the lowest level window for which independent data are analyzed: each distinct combination of angular resolution element, channel, and time yields a separate measurement and can count as an observation. Other useful windows are the survey \mathbf{s} , a complete collection of observations; and $\mathbf{h}(t)$, picking all objects in past history to the present as viewed from Earth. Still more are listed in Table 3.²¹

A collection S_y^x of windows can be fused to form another,

$$W_y^J = \bigcup_{x_i \in S_y^x} W_{x_i}^J, \quad (18)$$

²¹ Although not used in this work, windows can be random. These would be a point process on their own haystack. In fact, Wright et al. (2018) basically describe a survey window haystack (see also Djorgovski et al. 2013).

and if their selections are independent,

$$P_y^J(w_J) = 1 - \prod_{x_i \in S_y^x} (1 - P_{x_i}^J(w_J)). \quad (19)$$

The number of these subwindows is N_y^x . For example, a survey (S) is divided into N_S^0 observation windows (generically labeled O_i) from the set S_S^0 . The window resulting from fusing two subwindows x and y in this way can be written $x \cup y$. A joint window $x \cap y$ instead picks an event only if it is selected by both subwindows.

The last operation on windows is projection, working on a window or host and a fundamental variable ρ like time or frequency. The projection $x:\rho$ (or $K:\rho$) makes a selection according to the range of ρ covered by x or K (Figure 4). The projection of a window is given by

$$W_{x:\rho}^J = \bigcup_{\rho: W_x^J \cap W_\rho^J(\rho) \neq \emptyset} W_\rho^J(\rho) \quad (20)$$

and the projection of an object is

$$W_{K:\rho}^J(w_K) = \bigcup_{\rho: w_K \in W_\rho^K(\rho)} W_\rho^J(\rho) \quad (21)$$

with $P_{x:\rho}^J(w_J; \rho) = \max[P_x^J(w)P_\rho^J(w_J; \rho) | w, w_J \in W_\rho^J(\rho)]$ and $P_{K:\rho}^J(w_J; \rho, w_K) = \mathbb{I}[w_J \in W_{K:\rho}^J(w_K)]$. Projection is a versatile operation, letting us make selections according to a single shared quantity. For example, $\epsilon_{x:\nu|b}$ is the effective isotropic energy release of broadcast b at *any* time or polarization within the bandwidth of x ; $N_{c:t,g}^C$ is the number of communicative societies in galaxy g that are ever active during the lifespan of the specific society c . Projection is particularly useful when describing broadcasts that can run in and out of a window, like in the chord model later on (Appendix C).

By default, this series uses the ALL window (∞) selecting every event in a population when no other window is specified or implied.

3.2.2. Hosts

Specifying a host K in a selection restricts the population to its subpopulation, as defined by its point process: it draws J -type objects from $\Sigma_K^J(w_K)$. The host-based selection of objects is determined solely by their position on the tree—the object must be the direct ancestor or descendant of the host, or be the host itself. Objects are indexed, and the indices serve to denote their associated host selection as well. Most variables in this series use a host, but when none is otherwise specified, the cosmic host U is assumed.

3.2.3. Samples and Selection-relative Intensities

A full selection (x, L) , as a combination of a window and a host, selects J objects from the x -thinning of the point process $\Sigma_L^J(w_L)$ over the set W_x^J . The resulting point process, $\Sigma_{x,L}^J(w_L)$, is the random sample drawn by (x, L) . An object J is selected by (x, L) ($w_J \in \Sigma_{x,L}^J(w_L)$) if and only if:

1. It is within the random sample Σ_x^J ; that is, its parameter tuple w_J is within W_x^J and selected by the x -thinning.
2. It is a member of the subpopulation of its host ($w_J \in \Sigma_L^J(w_L)$), picked by the host selection L .

3. For each ancestor K of J up to and including L , $w_K \in \Sigma_{x,L}^K(w_L)$; so that each of these ancestors is *also* selected by the window.

This final condition is a somewhat tricky one, but an object can be missed because its ancestor was not sampled. Suppose we defined a window x that picked all galaxies and metasocieties in a patch of the sky, but only societies that are older than 1 million yr, and all their broadcasts. A radio broadcast from a society that is only 1000 yr old would not be selected by (x, U) even though $P_x^B(w_B) = 1$ for all w_B because its host society is excluded.

If L is a host for objects of type $K \in \mathcal{N}(J)$, the intensity relative to a selection (x, L) is

$$\Psi_{x,L}^J(w_J | w_L) = P_x^L(w_L) \int_{W^K} \Psi_L^K(w_K | w_L) \Psi_{x,K}^J(w_J | w_K) dw_K, \quad (22)$$

with

$$\Psi_{x,J}^J(w | w_J) = P_x^J(w) \Psi_J^J(w | w_J) = P_x^J(w) \delta(w_J - w). \quad (23)$$

Integrating gives us the expected number of J -type objects sampled by the selection

$$\langle N_{x,L}^J \rangle \equiv \langle N_{x,L}^J(w_L) \rangle = \int_{W^J} \Psi_{x,L}^J(w_J | w_L) dw_J. \quad (24)$$

The process of selection often introduces biases in object properties, so $\Psi_{x,L}^J$ is not usually proportional to Ψ_L^J . Generally, objects that are larger are overrepresented in a sample (see Section 3.8).

3.2.4. Simplifying Distributions through Marginalization, Rates, and Abundances

The full intensities often involve many parameters that are of no interest in a given problem. The distribution is simplified by marginalizing over these irrelevant quantities, treating them as marks to be ignored (Section 2.2). If a parameter tuple w_J can be divided into components that we are interested in, w_J^* , and those that we are not, w_J' , with $w_J = w_J^* \times w_J'$, then

$$\Psi_{x,K}^{J*}(w_J^* | w_K) = \int \Psi_{x,K}^J(w_J^* \times w_J' | w_K) dw_J'. \quad (25)$$

Of course, this equation applies to Ψ_K^J itself, which uses the ∞ window.

The most common of these marginalized intensities is the rate that objects form, a function over time, and possibly other fundamental variables like frequency. When ϑ_J is the time an object is born, the rate is

$$\check{I}_K^J(t, \nu | w_K) = \frac{d \langle N_{\nu,K}^J(\nu, w_K) \rangle}{d\vartheta_J}(\vartheta_J = t), \quad (26)$$

where $N_{\nu,K}^J(\nu, w_K)$ is the number of objects hosted by K and selected by a window picking all objects that cover the frequency ν . Broadcasts have limited bandwidth, but other types of objects (including societies, metasocieties, and galaxies) are not confined by frequency and have $N_{\nu,K}^J(\nu, w_K) = N_K^J(w_K)$. If the number of selected objects is proportional to the number of stars (e.g., more stars means more inhabited worlds), a convenient variable is the stellar rate,

$$I_K^J(t, \nu | w_K) = \check{I}_K^J(t, \nu | w_K) / \langle N_{t,K}^*(t, w_K) \rangle, \quad (27)$$

where $N_{t,K}^*(t, \mathbf{w}_K)$ is the number of stars extant at time t within the host K .

If $\Psi_{x,K}^J$ is marginalized over all parameters, then the resulting zero-dimensional distribution simply gives $\langle N_{x,K}^J \rangle$ as per Campbell's formula (Equation (4)). This itself can be regarded as a parameter, with $\langle N_{x,K}^J \rangle$ defined for some well-characterized window x scaling any other distribution. For objects that trace stars, the most useful of these scaling parameters is the stellar abundance

$$\Xi_K^J(t, \nu | \mathbf{w}_K) = \langle N_{t \cap \nu, K}^J(t, \nu, \mathbf{w}_K) \rangle / \langle N_{t,K}^*(t, \mathbf{w}_K) \rangle, \quad (28)$$

with $t \cap \nu$ picking all objects active at both time t and, for broadcasts, frequency ν . It is the mean number of objects per star that cover a point on a spectrogram.

3.3. Random Variables

3.3.1. Singleton Variables

A *singleton variable* κ_J is a random variable that describes a single object J .²² I assume that it is independent of all other objects, depending only on \mathbf{w}_J . Examples include the position of an object, luminosity, flux received at Earth, and the total number of photons collected in a telescope from that object in 1 pixel of a detector. The value of a singleton variable is not always determined by \mathbf{w}_J , only its distribution is. The number of photons counted even from a perfectly coherent laser beacon has a detector and shot noise, for example.

The basic notation for singleton variables is $\kappa_{x|J}$, where κ may be replaced with another symbol, usually a lowercase letter. In some cases, a random variable depends on how long we are observing, in what frequencies, at what polarizations, as is true for emission measured. The *quantity window* x sets the bounds of the integration for κ (Figure 5). Generally, an object with no activity or emission within this window is not selected by x ,

$$\kappa_{x|J}(\mathbf{w}_J) = 0 \Rightarrow \mathbf{w}_J \notin W_x^J, \quad (29)$$

although an object can be excluded for other reasons. When no quantity window is specified, the ALL window implicitly is used ($\kappa_J = \kappa_{\infty|J}$). The basic parameters for each object can be interpreted as special singleton variables that naturally use the ALL window; they are intrinsic properties, unaffected by observation.

A calculation with singleton variables over a population $\Sigma_{y,K}^J(\mathbf{w}_K)$ is performed by marking each point with the value of $\kappa_{x|J}(\mathbf{w}_J)$. As noted in Section 2.2, this is equivalent to considering a point process in $W^J \times \mathbb{R}$ for a real-valued random variable. As long as the marks are independent of any and all other points in the process, the intensity of this new process is

$$\frac{d^2 \langle N_{y,K}^J \rangle}{d\mathbf{w}_J d\kappa_{x|J}}(\mathbf{w}, \kappa | \mathbf{w}_K) = \Psi_{y,K}^J(\mathbf{w} | \mathbf{w}_K) \cdot \psi[\kappa_{x|J}](\kappa | \mathbf{w}). \quad (30)$$

²² Technically, for a random object J , $\kappa_J(\mathbf{w}_J)$ is a random function that outputs a random variable from the input \mathbf{w}_J (Section 3.1.3). The same caveat applies to aggregate variables with random hosts.

3.3.2. Aggregate Variables

An *aggregate variable* K is the sum of the singleton variables for a sample of objects drawn from a host K :

$$K_{x|y,K}^J(\mathbf{w}_K) \equiv \sum_{\mathbf{w}_J \in \Sigma_{y,K}^J(\mathbf{w}_K)} \kappa_{x|J}(\mathbf{w}_J), \quad (31)$$

where K may be replaced with another symbol, usually an uppercase letter. In addition to the quantity window x , a potentially distinct *object window* is combined with the host K to define a selection that picks out which of the objects are included in a sum. Each $K_{x|y,K}^J$ is treated using the marked point process for the corresponding $\kappa_{x|J}$.

To simplify notation, the quantity window may be omitted to indicate that it is the same as the object window: $K_{x,K}^J \equiv K_{x|y,K}^J$ (Table 5). Applying Equation (29) allows us to expand or contract the object window:

$$K_{x,K}^J = K_{x|y,K}^J \text{ if } (\forall \mathbf{w}_J \in W_x^J) P_x^J(\mathbf{w}_J) = P_y^J(\mathbf{w}_J). \quad (32)$$

Aggregate and singleton variables can be interconverted,

$$\kappa_{y|K} \leftrightarrow K_{x|y,K}^J, \quad (33)$$

where the object window of a singleton variable is always identical to its quantity window. These simply reflect different ways of viewing the quantity: the aggregate variable interprets it as a sum over a population, while the singleton variable interprets it as an intrinsic property of the host. A common example in extragalactic astronomy is the mass of a galaxy's stars—whether it is viewed as the collective property of a stellar population or a property of a single galaxy, it is still the same quantity. This lets us build up high-level aggregates recursively, like the total emission from all broadcasts in all societies in a galaxy.

Number variables can be interpreted as a special case of aggregate variables:

$$N_{y,K}^J(\mathbf{w}_K) = \sum_{\mathbf{w}_J \in \Sigma_{y,K}^J(\mathbf{w}_K)} 1 = |\Sigma_{y,K}^J(\mathbf{w}_K)|. \quad (34)$$

3.4. Selection-relative Probabilities, Means, and Variances: The Effects of Selection Bias

A distribution operation \odot takes a variable and returns a function or number derived from its probability distribution—it needs a random variable, and cannot stand on its own. These operations include the CDF, PDF, mean, and variance. In this section, I consider how these operations are affected by selection biases.

3.4.1. Selection-relative Probabilities

Let us say we want to know the probability of an event E_J describing some property of $\kappa_{x|J}$. Now, this probability is associated with a random object J and thus is actually a random function of \mathbf{w}_J . Our general assumption has been that the probability distribution of any random variable $\kappa_{x|J}$ is fully determined by \mathbf{w}_J , automatically implying independence between other such variables and the number of objects. This is the simple probability of the event, $P(E_J) \equiv P(E_J | \mathbf{w}_J)$.

In turn, the probability $P(E_J)$ is also a function of \mathbf{w}_J . However, any distribution for the J objects themselves allows

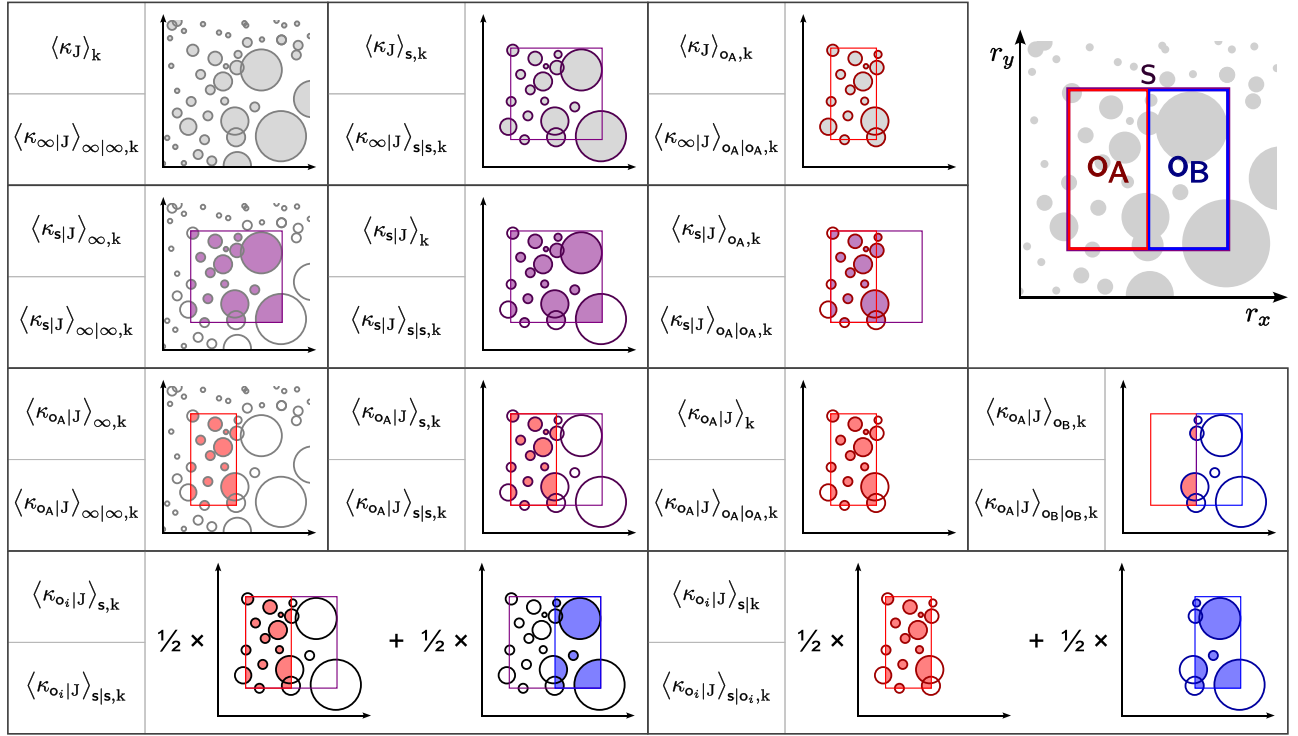


Figure 5. Comparison of different quantity windows and selections applied to a variable and an averaging operation. Here we have a cluster of J-type spherical objects in a host k , observed in the sky field as pictured at upper right. A survey S covers this field with two observations, O_A and O_B , each an aperture. Any object that touches a window is selected by it. The subplots depict different multiwindow selection-relative means, with simplified (above) and full (below) notation. The underlying variable κ is a quantity that is integrated over its quantity window (e.g., starlight from a galaxy that falls within the window), with the amount falling within its bounds indicated by the shading. The means average this quantity over only some of the objects, the ones outlined by rings. The multiwindow means on the bottom average over several observations; O_i is a placeholder for O_A and O_B . Note however that the means are actually ensemble averages, over all possible configurations of the objects, whereas only one realized configuration is shown. Other operations like the variance and PDF also have multiwindow selection-relative versions analogous to the means shown here.

Table 5
Simplifications in Notation for Variables and Operations

Singleton Variable		Aggregate Variable	
Short	Full	Short	Full
Variables			
κ_J	$\kappa_{\infty J}(\mathbf{w}_J)$	K_K^J	$K_{\infty K}^J(\mathbf{w}_K)$
		$K_{x,K}^J$	$K_{x K}^J(\mathbf{w}_K)$
		$K_{x K}^J$	$K_{x K}^J(\mathbf{w}_K)$
		K_x^J	$K_{x K}^J(\mathbf{w}_U)$
		$K_{x y}^J$	$K_{x y,U}^J(\mathbf{w}_U)$
Operations on Variables			
$\mathbb{O}[\kappa_{x J}]$	$\mathbb{O}[\kappa_{x J}]_{x J}$	$\mathbb{O}[K_{x y,K}^J]$	$\mathbb{O}[K_{x y,K}^J]_{y K}$
$\mathbb{O}[\kappa_{x J}]_K$	$\mathbb{O}[\kappa_{x J}]_{x K}$	$\mathbb{O}[K_{x y,K}^J]_L$	$\mathbb{O}[K_{x y,K}^J]_{y L}$
$\mathbb{O}[\kappa_{x J}]_y$	$\mathbb{O}[\kappa_{x J}]_{y J}$	$\mathbb{O}[K_{x y,K}^J]_z$	$\mathbb{O}[K_{x y,K}^J]_{z K}$
$\mathbb{O}[\kappa_{x J}]_{y,K}$	$\mathbb{O}[\kappa_{x J}]_{y K}$	$\mathbb{O}[K_{x y,K}^J]_{z,L}$	$\mathbb{O}[K_{x y,K}^J]_{z L}$
$\mathbb{O}[\kappa_{x J}]_{y K}$	$\mathbb{O}[\kappa_{x J}]_{y K}$	$\mathbb{O}[K_{x y,K}^J]_{z L}$	$\mathbb{O}[K_{x y,K}^J]_{z L}$

Note. The simplifications are applied from the inside out. Thus, $\langle \kappa_J \rangle_{y,K} \rightarrow \langle \kappa_{\infty|J} \rangle_{y,K} \rightarrow \langle \kappa_{\infty|J} \rangle_{y|y,K}$. When the host of a selection-relative operation (including a general multiwindow operation) is a random object, the result should be understood as a function of the host's properties.

us to generate a mixture distribution for $\kappa_{x|J}$, and then we can evaluate the probability of E_J for this mixture. This includes distributions modified by selection windows. When the value of a random variable $\kappa_{x|J}$ for a single object is independent of

all other objects, the selection-relative probability is thus defined as

$$P_{y,K}(E_J) \equiv \frac{1}{\langle N_{y,K}^J \rangle} \int_{\mathbf{w}_y^J} \Psi_{y,K}^J(\mathbf{w}_J | \mathbf{w}_K) P(E_J | \mathbf{w}_J) d\mathbf{w}_J. \quad (35)$$

This is the probability that E_J occurs for a J-type object fairly drawn from the random sample $\Sigma_{y,K}^J(\mathbf{w}_K)$. It is an ensemble probability, averaging over all possible selections, not just for a realized sample. It reflects the bias of a selection.

An archetypal event is $\kappa_{x|J}$ being within some specified range. The previous equation immediately yields a selection-relative CDF, with $F[\kappa_{x|J}]_{y,K}(x) \equiv P_{y,K}(\kappa_{x|J} \leq x)$. Differentiating then gives the selection-relative PDF,

$$\begin{aligned} & \psi[\kappa_{x|J}]_{y,K}(\kappa) \\ & \equiv \frac{1}{\langle N_{y,K}^J \rangle} \int_{\mathbf{w}_y^J} \Psi_{y,K}^J(\mathbf{w}_J | \mathbf{w}_K) \cdot \psi[\kappa_{x|J}](\kappa | \mathbf{w}_J) d\mathbf{w}_J, \end{aligned} \quad (36)$$

where $\psi[\kappa_{x|J}]$ is the (unbiased) simple probability distribution for $\kappa_{x|J}$ determined by \mathbf{w}_J . An analogous equation can be written for aggregate variables.

Note that the objects sampled by a selection are random. Thus, selection-relative probabilities and operations are used on things describing random objects, although the selection itself may have a realized host.²³

²³ If for some reason a selection-relative operation were applied to a variable describing a realized object, the selection would drop out, and it would become a simple operation described in the next subsection (e.g., $\psi[\kappa_{x|J}]_{y,K} = \psi[\kappa_{x|J}]$).

3.4.2. A Notational Convention for Simple and Selection-relative Operations

The simple probabilities are selection-relative probabilities where the self-point process is being used. This rule extends to all distribution operations on the probability

$$\begin{aligned}\mathbb{O}[\kappa_{x|J}] &= \mathbb{O}[\kappa_{x|J}|\mathbf{w}_J] = \mathbb{O}[\kappa_{x|J}]_{x,J} \\ \mathbb{O}[K_{x|y,K}^J] &= \mathbb{O}[K_{x|y,K}^J|\mathbf{w}_K] = \mathbb{O}[K_{x|y,K}^J]_{y,K},\end{aligned}\quad (37)$$

including mean, variance, and order statistics.

A similar simplification is used in this series for windows—a selection-relative distribution operation missing its window inherits it from the variable’s object window (which is the quantity window too for singleton variables):

$$\mathbb{O}[\kappa_{x|J}]_L \equiv \mathbb{O}[\kappa_{x|J}]_{x,L} \text{ and } \mathbb{O}[K_{x|y,K}^J]_L \equiv \mathbb{O}[K_{x|y,K}^J]_{y,L}, \quad (38)$$

including mean and variance (Table 5). Contrast with the convention for variables, which stand on their own and have no other variable to draw a window from. This convention is motivated by practicality: in most windows we use, the vast majority of objects do not contribute to aggregate quantities (Equation (29)) and do not concern us. When we want to know the mean broadcast fluence over an observation, we do not want to include those that were only visible during the Jurassic, for instance.²⁴ Remember, if the variable itself does not specify a window, it uses the ALL (∞) window (Section 3.2.1), which the mean then inherits ($\mathbb{O}[\kappa_{x|J}]_K = \mathbb{O}[\kappa_{x|J}]_{\infty,K}$). This applies to the intrinsic quantities describing objects like their lifespan and position.

3.4.3. Simple Means for a Single Object

The most straightforward averaging operation is the simple mean. Say we have a singleton quantity $\kappa_{x|J}$ associated with a random object of type J with a parameter tuple \mathbf{w}_J . Its simple mean is:

$$\langle \kappa_{x|J} \rangle \equiv \langle \kappa_{x|J}(\mathbf{w}_J) | \mathbf{w}_J \rangle = \int \kappa \cdot \psi[\kappa_{x|J}](\kappa | \mathbf{w}_J) d\kappa. \quad (39)$$

This is a function of \mathbf{w}_J . The simple mean of an aggregate variable is, with the help of Equation (33),

$$\begin{aligned}\langle K_{x|y,K}^J \rangle &\equiv \langle K_{x|y,K}^J(\mathbf{w}_K) | \mathbf{w}_K \rangle \\ &= \int K \cdot \psi[K_{x|y,K}^J](K | \mathbf{w}_K) dK.\end{aligned}\quad (40)$$

Recall that \mathbf{w}_K does *not* determine the subpopulation of K ’s descendants, only its statistical properties. The mean is therefore an average over all realizations of this subpopulation, conditionalized on the properties of the host. This sampling variance within the host is regarded as noise.

The average is found by applying Campbell’s formula (Equation (4)) to the intensity of the marked point process:

$$\langle K_{x|y,K}^J \rangle = \int_{\mathbf{w}_y^J} \langle \kappa_{x|J} \rangle \Psi_{y,K}^J(\mathbf{w}_J | \mathbf{w}_K) d\mathbf{w}_J. \quad (41)$$

The integral for $\langle \kappa_{x|J} \rangle$ in Equation (39) is just a special case of this, summing over $\Sigma_{y,K}^J(\mathbf{w}_J)$.

²⁴ If we *do* want to include these objects, we apply the ALL window operator, $\mathbb{O}[X]_{\infty,K}$ (compare Figure 5).

3.4.4. Selection-relative Means

But what if we want to average, not just over the noise in a random variable, but over different possible draws of its associated object? This is what a selection-relative mean does (Figure 5). For a selection (Z, L), the selection-relative mean is

$$\begin{aligned}\langle \kappa_{x|J} \rangle_{Z,L} &\equiv \langle \kappa_{x|J}(\mathbf{w}_J) | \mathbf{w}_J \in \Sigma_{Z,L}^J(\mathbf{w}_L) \rangle \\ \langle K_{x|y,K}^J \rangle_{Z,L} &\equiv \langle K_{x|y,K}^J(\mathbf{w}_K) | \mathbf{w}_K \in \Sigma_{Z,L}^K(\mathbf{w}_L) \rangle.\end{aligned}\quad (42)$$

Means with respect to realized hosts have analogous definitions. Unlike the simple means, the selection-relative means averages over all the different possible parameter tuples for the object being described by the random variable, weighted by the probability of it being a part of the sample drawn by the selection. It is an *ensemble* average, the mean that we would get if we repeated the selection again and again.

To give a concrete example, suppose we had a society c in a galaxy g , and we observe it over a window \mathbf{o} . The total energy received from all broadcasts in that society over \mathbf{o} is $\mathfrak{E}_{\mathbf{o},c}^B$. Then $\langle \mathfrak{E}_{\mathbf{o},c}^B \rangle$ is the mean value of the energy we receive from the broadcasts in that society, or more precisely, the average for an ensemble of societies each with an identical parameter tuple \mathbf{w}_c . However, $\langle \mathfrak{E}_{\mathbf{o},C}^B \rangle_g$ is the mean value over all the possible societies in the galaxy, roughly the value of $\mathfrak{E}_{\mathbf{o},C}^B$ expected from a typical society.

The selection-relative mean is found using the selection-relative (biased) PDF (Equation (36)). Assuming all the variables are mutually independent, Equation (39) gives us

$$\langle \kappa_{x|J} \rangle_{Z,L} = \frac{1}{\langle N_{Z,L}^J \rangle} \int_{\mathbf{w}_Z^J} \langle \kappa_{x|J} \rangle \Psi_{Z,L}^J(\mathbf{w}_J | \mathbf{w}_L) d\mathbf{w}_J, \quad (43)$$

with the selection-relative mean of an aggregate variable found with an analogous integral (Equation (33)). A consequence of Campbell’s formula is that

$$\langle K_{x|y,K}^J \rangle = \left\langle \sum_{\mathbf{w}_J \in \Sigma_{y,K}^J(\mathbf{w}_K)} \kappa_{x|J}(\mathbf{w}_J) \right\rangle = \langle N_{y,K}^J \rangle \langle \kappa_{x|J} \rangle_{y,K}. \quad (44)$$

Once completed, the averages may themselves be averaged, most naturally when the inner average’s selection host is a random object.²⁵ As a random variable, this variable inherits the object window and host from the averaging operator itself: $\langle X_J \rangle_{y,K} \rightarrow \kappa'_{y|K}(\mathbf{w}_K)$.

Note that, generally,

$$\langle \langle \kappa_{x|J} \rangle_{y,K} \rangle_{Z,L} \neq \langle \kappa_{x|J} \rangle_{Z,L}, \quad (45)$$

and similarly for aggregate variables. That is because the left-hand side averages over subselections before averaging over the outermost selection. Some subselections may have more J -type objects than others, but these are all weighted equally on the left-handed side, despite their different population sizes. By evenly sampling the K -type host objects, the left-hand side fails to evenly sample the J -type objects like the right-hand side. This effect is related to Simpson’s paradox (Haigh 2013).

²⁵ If the operation’s host is a realized object k , then the selection-relative mean (or variance) has a fixed value determined by \mathbf{w}_k (i.e., it is a degenerate random variable); whereas, if the operation’s host is random object K , then it is a function that yields a single value for each \mathbf{w}_K , or equivalently, the degenerate random variable that always attains that value.

However, it is true that

$$\langle \langle \kappa_{x|J} \rangle \rangle_{z,L} = \langle \kappa_{x|J} \rangle_{z,L} \quad (46)$$

because the inner mean's selection necessarily has a population size of 1—this is just the law of total expectation, and an expression of Campbell's formula for marked point processes.

3.4.5. Simple and Selection-relative Variances

The simple variance of a variable is found using the distribution of that variable that comes from its tuple. Selection-relative variances then average over a distribution of random objects. Their definition follows from that of a general conditional variance, $\mathbb{V}[X|Y] = \langle X^2|Y \rangle - \langle X|Y \rangle^2$ (Wasserman 2004; Bas 2019). Thus, $\mathbb{V}[X]_{x,K} = \langle X^2 \rangle_{x,K} - \langle X \rangle_{x,K}^2$.

The law of total variance is an important tool in evaluating both types of variances. By first conditionalizing on $N_{y,K}^J(\mathbf{w}_K)$, we derive the simple variance of $K_{x|y,K}^J(\mathbf{w}_K)$:

$$\mathbb{V}[K_{x|y,K}^J] = \langle \kappa_{x|J} \rangle_{y,K}^2 \mathbb{V}[N_{y,K}^J] + \langle N_{y,K}^J \rangle \mathbb{V}[\kappa_{x|J}]_{y,K}, \quad (47)$$

as long as all of the variables are independent. When $N_{y,K}^J(\mathbf{w}_K)$ for a fixed \mathbf{w}_K is also Poissonian, the variance reduces to

$$\langle N_{y,K}^J \rangle \langle \kappa_{x|J}^2 \rangle_{y,K} = \int_{\mathbf{w}_J} \langle \kappa_{x|J}^2 \rangle \Psi_{y,K}^J(\mathbf{w}_J | \mathbf{w}_K) d\mathbf{w}_J,$$

equivalent to Campbell's second formula (Equation (10)). The law also gives us, for selection-relative variances,

$$\mathbb{V}[K_{x|y,K}^J]_{z,L} = \mathbb{V}[\langle K_{x|y,K}^J \rangle]_{z,L} + \langle \mathbb{V}[K_{x|y,K}^J] \rangle_{z,L}. \quad (48)$$

3.5. Multiwindow Operations

An observing program is usually a collection of many observations: a survey \mathbf{s} might take several pointings π_i , each in turn involving thousands of distinct observations \mathbf{o}_j . The problem is that the statistical properties of a sample may depend on which subwindow we are considering. Differences between subwindows also introduce still more variance. When mapping a spiral galaxy, for instance, the stars are far more dense in the core than the outer disk, increasing the variance in star counts because the underlying density changes. We need to be able to ask questions like, What is the expected number of stars per observation averaged over all observations in the survey? and What is the maximum broadcast fluence during an observation among all the observations we do?

Just as the selection-relative operations have an object window that picks which objects are included in the operation, the multiwindow operations have a quantity window, which defines which subwindows are considered. The most general operation is written as $\mathbb{O}[X]_{z|y,K}$, where z is its object window, and (y, K) is its selection.

For our purposes, there are a fixed number of subwindows with specified properties. I define the PDF as a mixture of the PDFs for the individual windows:

$$\psi[\kappa_{x|J}]_{z|y,K}(\kappa | \mathbf{w}_K) \equiv \frac{1}{N_z^x} \sum_{i=1}^{N_z^x} \psi[\kappa_{x|J}]_{y,K}(\kappa | \mathbf{w}_K). \quad (49)$$

The general multiwindow mean is

$$\langle \kappa_{x|J} \rangle_{z|y,K} \equiv \langle \langle \kappa_{x|J} \rangle_{y,K} | \mathbf{x}_i \in \mathbf{S}_z^x \rangle. \quad (50)$$

We simply take the mean over the given selection for each object window, then take the average over the object windows

(Figure 5). The general multiwindow variance is, by the law of total variance,

$$\begin{aligned} \mathbb{V}[\kappa_{x|J}]_{z|y,K} &\equiv \langle \kappa_{x|J}^2 \rangle_{z|y,K} - \langle \kappa_{x|J} \rangle_{z|y,K}^2 \\ &= \langle \mathbb{V}[\kappa_{x|J}]_{y,K} | \mathbf{x}_i \in \mathbf{S}_z^x \rangle + \mathbb{V}[\langle \kappa_{x|J} \rangle_{y,K} | \mathbf{x}_i \in \mathbf{S}_z^x]. \end{aligned} \quad (51)$$

The second term in the final equality is the variance due to the changes in the population between \mathbf{x}_i windows. Finally, the general multiwindow maximum is

$$\max[\kappa_{x|J}]_{z|y,K} \equiv \max[\max[\kappa_{x|J}]_{y,K} | \mathbf{x}_i \in \mathbf{S}_z^x], \quad (52)$$

and likewise for the minimum. Thus, the mean energy fluence from a broadcast during a typical observation in a survey \mathbf{s} among the broadcasts the survey selects from a galaxy \mathbf{g} is written $\langle h_{o|B} \rangle_{s|s,g}$, for instance.

The populations in different windows are not generally independent. Adjacent windows can cover the same objects: the sidelobes of one pointing can intersect the primary beam of another for instance. Sometimes the objects themselves are extended, crossing into different windows, as when a frequency-drifting narrowband broadcast cuts across many channels. The greatest effect of this cross correlation is on the maxima and minima because they reduce the effective number of independent samples. If the subwindows are very broad, the variance can be reduced as well because the same objects are covered in all pointings, averaging out each measurement.

When working with multiwindow operations, there is a distinction between general unspecified windows and specific windows, similar to that between random objects and realized objects. We can assign a unique label to a unique window, like \mathbf{o}_1 or \mathbf{o}_2 , and apply it to a variable. This fixes a unique subwindow; the operation's object window becomes irrelevant because the subwindow is one of a kind. In other cases, we really do want to consider the full gamut of subwindows in a collection, and to emphasize this, we might use a subwindow label like \mathbf{o}_i , where i is a free variable. Most often, I assume that the statistical properties of the population do not change appreciably from window to window. Then, the multiwindow selection-relative means and variances simply collapse to the selection-relative means and variances, and the multiwindow maxima are a simple application of extreme value theory (Appendix A). Only then can we elide the distinction between unspecified and specified windows: any observation is just an observation \mathbf{o} .

The multiwindow operations come with their own notational simplifications. The quantity window can be inherited from the operation's selection but not the variable, while an operation labeled *solely* by a window uses it as a quantity window:

$$\begin{aligned} \mathbb{O}[\kappa_{x|J}]_{y,K} &\equiv \mathbb{O}[\kappa_{x|J}]_{y|y,K} \\ \mathbb{O}[\kappa_{x|J}]_y &\equiv \mathbb{O}[\kappa_{x|J}]_{y|y,K}. \end{aligned} \quad (53)$$

Table 5 summarizes all of the notation rules, and Figure 5 provides an illustrated example. As before, the object window of an operation is inherited from the variable, even when a separate quantity window is specified:

$$\mathbb{O}[\kappa_{x|J}]_{y|K} \equiv \mathbb{O}[\kappa_{x|J}]_{y|y,K}. \quad (54)$$

3.6. Regularization of Variables with Broad Distributions

Like other astrophysical entities, ETIs and technosignatures may be described by quantities with heavy-tailed distributions. The radiated energy is widely thought to be a possible example (Drake

et al. 1973); others may include the lifespan of societies or transmitters, or the number of societies per metasociety. Plausible distributions like shallow power laws can have variances and means dominated by very rare events with extreme properties. These objects may be too uncommon to detect in any practical selection, however, and thus throw off estimates for the properties of a typical sample.

When needed, this series estimates mean and variance by trimming values of a random variable that are far outside the usual range in a typical sample. This regularization is done by imposing cutoff values κ^L and κ^H based on extreme value statistics (Appendix A). Starting from a random variable $\kappa_{x|J}$ and trimming with respect to a selection (y, K) , the regularized variable $\kappa_{x|J}^{[y,K]}(\mathbf{w}_J; \mathbf{w}_K)$ has a truncated distribution,

$$\psi[\kappa_{x|J}^{[y,K]}](\kappa) \equiv \frac{\psi[\kappa_{x|J}](\kappa) \cdot \mathbb{I}[\kappa^L \leq \kappa \leq \kappa^H]}{P_{y,K}(\kappa^L \leq \kappa_{x|J} \leq \kappa^H)}. \quad (55)$$

Equation (55) also regularizes aggregate variables, as they can be treated as just a special case of singleton variables (Equation (33)). Because aggregate variables are sums, the calculations are more complicated. If the individual contributions have a narrow enough distribution that their full range is expected to be well sampled in each aggregate, any large fluctuations in the sum must be due to fluctuations in the number of objects included: $K_{x,K}^{J:[y,L]} \approx N_{x,K}^{J:[y,L]} \langle \kappa_{x|J} \rangle_{x,K}$. These fluctuations may be relatively small and the regularization has essentially no effect. If the individual contributions have a very broad distribution—when we are calculating the aggregate luminosity of objects with a shallow power-law distribution, for example—the sum is almost entirely dominated by the largest contribution in the sample and $K_{x,K}^{J:[y,L]} \approx \max[\kappa_{x|J}]_{y,L}$.

The cutoff is a statistical quantity, however, and any actual sample may include extreme outliers. Thus, this regularization will underestimate the sample mean and variance some of the time, possibly by a large degree in rare instances.

3.7. The Discreteness Criterion

In order to make a detection of a broadcast in a survey \mathbf{s} , at least one must be sampled during a program,²⁶ and a null result should occur whenever zero broadcasts are sampled. It is possible that our observation was too short, too narrowband, or had too small a field to intercept rare broadcasts even when they are extremely bright. This happens outside of SETI too: a supernova or a quasar episode can boost a galaxy's luminosity far above its typical value for a short time. If a galaxy's measured luminosity is fainter than a supernova or a quasar, that just means one was not active during the observation, not that they never are found there. Thus, a null result is compatible with any models with a high enough probability of zero broadcasts being detected.

The discreteness criterion states that we should only expect a detection in survey \mathbf{s} with models having

$$P(N_s^B = 0) \leq \bar{P} \quad (56)$$

for a suitably conservative false negative threshold \bar{P} near 0. Only models that fulfill this criterion are likely to be constrained by observations. To order of magnitude, we expect a nonempty sample only if the mean number of objects is $\gtrsim 1$.²⁷

²⁶ The converse is not always true—a sampled broadcast may be too faint to detect, or confused with many others.

²⁷ This follows from Markov's inequality, which gives $P(N_s^J \geq 1) \leq \langle N_s^J \rangle$ for the nonnegative N_s^J (Wasserman 2004).

However, no broadcasts will be detected if the sample is missing any of the ancestor nodes on the tree—societies, metasocieties, or galaxies. Equation (56) implies the often-simpler formula

$$\langle N_s^K \rangle \geq 1 - \bar{P} \approx 1 \text{ if } K \in \mathfrak{N}(B) \text{ or } K = B. \quad (57)$$

3.8. An Example: Lifespan Bias in Time-limited Selection

Long-lived objects are disproportionately represented in a field once a population has reached equilibrium. This phenomenon serves as an example of how to make selection-relative calculations.

Suppose we have objects of type J in a host K . How is the probability distribution of object lifespan affected by a time-limited selection? To start, let us suppose that each such object forms at time ϑ_J and exists for duration τ_J without interruption. The distribution function for these objects has the form

$$\Psi_K^J(\mathbf{w}_J | \mathbf{w}_K) = \frac{d^3 \langle N_K^J \rangle}{d\tau_J d\vartheta_J d\mathbf{w}_J}(\mathbf{w}_J | \mathbf{w}_K), \quad (58)$$

where \mathbf{w}_J' stands for a tuple of other, irrelevant parameters to marginalize over. Now consider a window \mathbf{x} that selects all objects in K active anytime during the contiguous time interval $[\Theta_x, \Theta_x + T_x]$. Clearly, an object in this population is selected if it either formed before Θ_x and lives long enough to overlap the window ($\Theta_x - \tau_J \leq \vartheta_J \leq \Theta_x$) or if it forms within the window ($\Theta_x \leq \vartheta_J \leq \Theta_x + T_x$). Thus, if the objects form at a constant rate \check{I}_K^J (Equation (27)),

$$\begin{aligned} \frac{d \langle N_{x,K}^J \rangle}{d\tau_J}(\tau | \mathbf{w}_K) &= \int_{\Theta_x - \tau}^{\Theta_x + T_x} \Psi_K^J(\mathbf{w}_J | \mathbf{w}_K) d\mathbf{w}_J' d\vartheta_J \\ &= (\tau + T_x) \cdot \check{I}_K^J(\mathbf{w}_K) \cdot \psi[\tau_J]_K(\tau) \end{aligned} \quad (59)$$

from Equation (24), where $\psi[\tau_J]_K$ is the unbiased lifespan distribution. It follows immediately that $\langle N_{x,K}^J \rangle = (\langle \tau_J \rangle_K + T_x) \cdot \check{I}_K^J(\mathbf{w}_K)$, giving us the \mathbf{x} -biased duration distribution (Equation (36))

$$\psi[\tau_J]_{x,K}(\tau) = \frac{\tau + T_x}{\langle \tau_J \rangle_K + T_x} \psi[\tau_J]_K(\tau). \quad (60)$$

Finally, we have (Equation (43))

$$\langle f(\tau_J) \rangle_{x,K} = \frac{\langle \tau_J f(\tau_J) \rangle_K + T_x \langle f(\tau_J) \rangle_K}{\langle \tau_J \rangle_K + T_x}. \quad (61)$$

These formulae assume that the rate has been constant infinitely far into the past, at least in that each object's lifespan is shorter than the age of K . Otherwise, an additional cutoff is needed for the integration in Equation (59), as is well known in the field of stellar populations (e.g., Miller & Scalo 1979). In the limit that all the objects are more short-lived than the host's age at the start of \mathbf{x} , $\Theta_x - \vartheta_K$, Equation (61) still applies. On the other hand, if most all the J objects live longer than $\Theta_x - \vartheta_K$, it can be shown the bias disappears: almost every object that has ever been born is still around, so we get a fair sample of them.

When the observational window is short compared to typical lifespans, Equation (61) effectively weights the biased average by another power of τ_J . Every biased moment depends on the next higher unbiased moment, save the zeroth where a

cancellation occurs. For the mean lifespan in particular, $\langle \tau_j \rangle_{x,K} = \langle \tau_j \rangle_K + \mathbb{V}[\tau_j]_K / (\langle \tau_j \rangle_K + T_x) \geq \langle \tau_j \rangle_K$. Hence, one can replicate the result in Kipping et al. (2020) that $\langle \tau_j \rangle_{x,K} \approx 2 \langle \tau_j \rangle_K$ for exponentially distributed τ_j . Or, one can show that the lifespan's selection-relative mean is ill-behaved for a power-law distribution $\psi[\tau_j]_K \propto \tau_j^{-\alpha}$ extending to infinity unless $\alpha > 3$.

Lifespan bias can be thought of in two ways (Figure 6). In real spacetime, bigger objects are more likely to cross into a window with a fixed duration, having a larger cross section. In the haystack, the window itself becomes wider for longer lifespans, and thus, has a longer reach for these objects. A similar bias applies to band-limited selection by frequency: wideband broadcasts are more likely to be intercepted, a significant fact in the chord model presented later.

Figure 6 helps illustrate the different types of means. Each dot represents a specific, realized object; $\langle \tau_j \rangle$ is the simple mean of the lifespan for that object j , which is entirely determined by the dot's position in the haystack. Then $\langle \tau_j \rangle = \langle \tau_j \rangle_I(\mathbf{w}_j)$ is a *function* ranging over the entire space that returns the average lifespan for an object given its parameters—it returns the τ_j coordinate in the right-hand panel. A realized host k has an accompanying distribution of objects over the haystack, which we can average over: $\langle \tau_j \rangle_k$ averages over the entire haystack, while $\langle \tau_j \rangle_{x,k}$ averages only over the blue-shaded region. One such realization of the distribution is shown in Figure 6. Likewise, for a random host K , the distribution itself is a function of \mathbf{w}_K , as are $\langle \tau_j \rangle_K(\mathbf{w}_K)$ and $\langle \tau_j \rangle_{x,K}(\mathbf{w}_K)$.

4. Assumptions about Randomness

Technosignatures are the result of complex biological, social, and technological phenomena, with a panoply of factors shaping their populations. The full interplay responsible for broadcast characteristics may be analytically intractable. Nonetheless, treating broadcasts as the result of a hierarchy of conditionally independent point processes illuminates our understanding of the ETI populations.

4.1. Random ETIs

Random variables are a very general concept (Section 2.1), and can describe deterministic events. Still, the rest of the paper views broadcasts and other objects as unpredictable. This does not mean that the intelligences behind them have no specific motivations. Random processes can be used to describe human behavior by focusing on statistical trends rather than the unpredictable specific outcomes resulting from complex motivations (e.g., Jusup et al. 2022); the same seems plausible for ETIs.

The most basic properties of metasocieties and societies are where they are, when they start, and how long they last. The first two of these likely involve stochastic elements. Potential habitats are scattered randomly through the galaxy, and the timing of ETI evolution likely involves some contingency, although subsequent interstellar expansion might not.

Broadcasts can have predictable structure, though. Predictability can be a byproduct of how a transmitter operates. A narrow-beam transmitter on a rotating world sweeps past Earth at regular intervals, for example (Gray & Ellingsen 2002). ETIs could even exploit periodicity to encode information or make the artificial nature of their broadcasts obvious (as in Borra 2012; Harp et al. 2018).

However, although broadcasts may have predictable features, their properties are not fully determined—otherwise, we would simply aim our telescopes where and when we expect a broadcast instead of sifting through the haystack. Broadcasts plausibly turn on and off at unpredictable times, the result of unknowable social factors. Randomness in the host galaxy also induces randomness in broadcasts. Each host star has random light travel delay times, parallaxes, and Doppler shifts, and sits in a turbulent interstellar medium with random dispersion and scintillation. A faint periodic broadcast might only be observable during rare moments when scintillation magnifies it to a detectable level (Cordes et al. 1997).

The real issue is whether the treatment of broadcasts as independent isolated bursts of radiation is inadequate, and thus, the haystack is defined according to the wrong parameters. Introducing dependence is one way to address this: in a periodic train, knowing two subsequent arrivals allows us to predict the others. Often there is a natural parameterization in which we can treat the entire complex of signals as a single broadcast, however, like defining a periodic train by its period and phase. Although the specific distributions used in this paper would not apply, the general ideas of the analysis still do.

4.2. Independence

I also assume the properties of two sibling objects are independent of each other, conditionalized on the shared parent's parameter tuple. Likewise, the number of children objects in a parent is conditionally independent of each child object's parameters. Thus, each broadcast is assumed to be independent of every other, except in the sense they depend on a shared ancestor (society, metasociety, galaxy, or the universe); societies are independent of each other, aside from their shared ancestors, and so on.

Independence is by no means obvious. On the broadcast level, although beacons and noncommunicative broadcasts are plausibly completely independent, intragalactic communications are part of a web of messages, providing a natural mechanism for coordination. Broadcasts may be more common if there are more potential receiving societies, and some may even be direct replies to others. Metasocieties can converge to a common broadcast protocol, like designated frequency channels. Dependence could extend to other object types—if interstellar expansion is rampant, the existence of a society around one star strongly suggests the existence of them around nearby stars, and so on.

The framework handles all such seeming dependence by moving the shared properties up the tree. Instead of viewing sibling objects as mutually dependent, we instead postulate that the dependent quantities are actually properties of the ancestor itself. Thus, the common properties all follow from the ancestral tuple, which governs the populations within it. *Conditionalized* on that tuple, the siblings are independent of each other. The difference is that the population does not respond to its own fluctuations.

Consider what if each galaxy has exactly one society arising at a random time since the Big Bang, and the society's lifespan is always 100 yr? Without knowing when the society existed, a single broadcast from that society could occur at any time in cosmic history. Once we know when one broadcast occurred, however, any others had to have been made within 100 yr because there was no broadcasting society at other times—a failure of independence. However, that is only because the

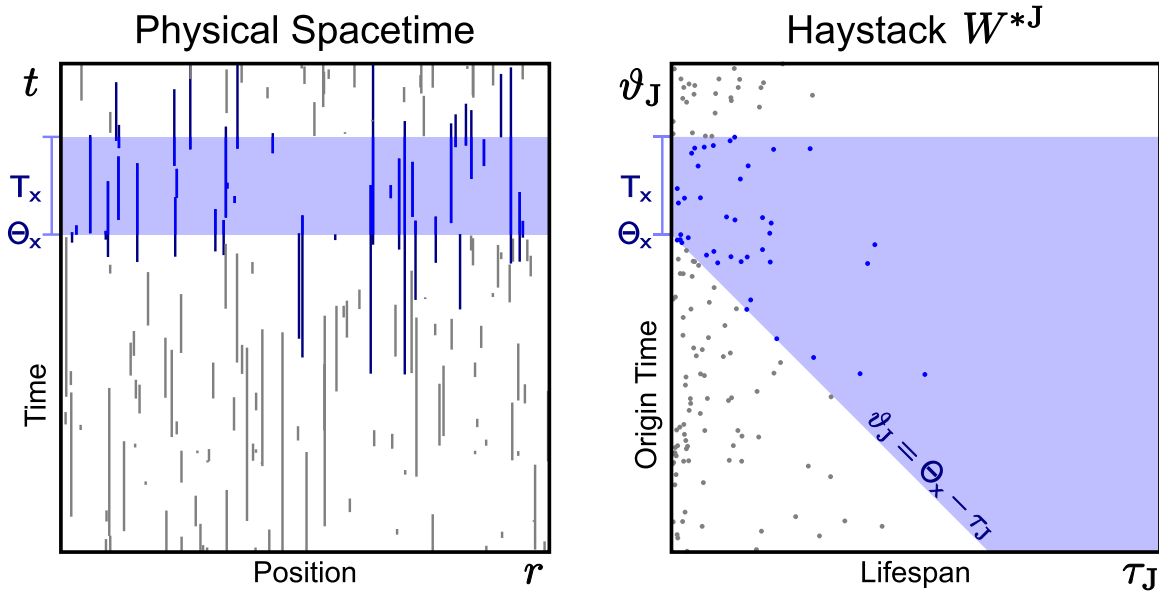


Figure 6. Lifespan bias by a window x in a population of objects with exponentially distributed lifespans. Selected objects are in blue (bright blue within the window). On the left, the objects are scattered in spacetime, and their lifespans are intervals in time. Those with longer τ_J are more likely to reach the window, which covers a fixed span of real time (blue shading). On the right, the objects are represented by points in a simplified haystack (spatial position not included). The window covers a bigger ϑ_J range as τ_J increases.

broadcast tells us when the society appeared. If we already knew that, learning when one broadcast happens does not necessarily give us any new information—the broadcast times can be independent and conditional on the societal parameters.

The model can be altered to include different levels to allow for still further degrees of dependence, but it nonetheless assumes that everything can be grouped into discrete objects sorted into hierarchical levels. Perhaps ETI distributions are more like scale-free fractals with a continuous range of dependence. This would necessitate a different kind of framework.

Just because the properties of one object are conditionally independent of another’s does not mean that the parameters of a single object are mutually independent. Massive stars are younger on average; metasocieties encompassing more worlds could live longer; gregarious societies could transmit broadcasts at a higher rate and make them brighter on average, and so on. This just means that the joint distribution is not separable into the product of the PDFs of the variables, like when all the objects fall along a line or surface in the haystack. It prevents us from modeling the aggregate emission with compound Poisson statistics. The distribution can still be calculated, however, and we still find the mean with Campbell’s formula (Equation (4)). If the societal population is Poissonian, Equation (10) allows us to find the variance.

5. The Astrophysical Context of ETIs: The Universe, Galaxies, and Stars

5.1. The Universe

The universe is the root node of the tree (Figure 1), the object that contains all the other objects as descendants. Even the universe has a parameter tuple, w_U , and a haystack, the space of all such tuples W^U (Table 6 lists relevant notation). Cosmological parameters like the Hubble constant can be used to define this haystack, but in this work, I assume these are set to fixed values. Instead, the model parameters describe the

descendant populations of ETIs and their broadcasts. For instance, suppose that we thought that in every transmitter population, the luminosity distribution follows a power-law distribution, all with the same slope γ . We do not know what the slope is, however. Then we can consider γ_U as a cosmic-level parameter, one describing all ETI populations in the Universe. In a given model universe u , it takes on a value γ_u , which all the lower-level populations inherit, and then we make calculations for each γ_u variate to compare predictions.

As discussed in Section 3.1.2, no universe distribution is considered here, only the consequences of different w_U for observed populations.

5.2. Galaxies as Domains for ETIs

ETIs can coordinate their properties on a large scale through replication and communication, but there must be a limit to how far these processes can operate. We can imagine the Universe ends up broken up into a patchwork of regions over which coordination succeeded, much like the magnetic domains in a piece of iron (see Olson 2015). In this series, a domain is an astrophysical region so large and isolated that ETIs cannot spread outside of it, completely confining any metasocieties. Depending on how difficult space travel is, a domain might be as small as a planetary body or as large as the entire Hubble volume. Every domain has sites where individual societies can reside. This work treats galaxies as the natural scale for domains.²⁸

Although they are home to entire populations of discrete stars and worlds, galaxies themselves form a cosmic population

²⁸ If intergalactic travel or communication is allowed, galactic metasocieties could fuse into an intergalactic metasociety on the domain scale. If cohesive galactic submetasocieties persist, then these would be the children of the top-level intergalactic metasocieties, instead of being the parents. We could also consider galaxies analogously to stars in Section 5.3.

Table 6
Shared Notation for Objects

Quantity	J-type	Universe	Galaxy	Metasociety	Society	Broadcast	Star
Random object index	J	U	G	M	C	B	★
Realized object index	j	u	g	m	c	b	◇
Parameter tuple	$\mathbf{w}_J, \mathbf{w}_j$	$\mathbf{w}_U, \mathbf{w}_u$	$\mathbf{w}_G, \mathbf{w}_g$	$\mathbf{w}_M, \mathbf{w}_m$	$\mathbf{w}_C, \mathbf{w}_c$	$\mathbf{w}_B, \mathbf{w}_b$	$\mathbf{w}_\star, \mathbf{w}_\diamond$
Position	$\mathbf{r}_J, \mathbf{r}_j$...	$\mathbf{r}_G, \mathbf{r}_g$	$\mathbf{r}_M, \mathbf{r}_m$	$\mathbf{r}_C, \mathbf{r}_c$	$\mathbf{r}_B, \mathbf{r}_b$	$\mathbf{r}_\star, \mathbf{r}_\diamond$
Size (volume)	v_J, v_j	v_U, v_u	v_G, v_g	v_M, v_m
Starting time	ϑ_J, ϑ_j	ϑ_U, ϑ_u	ϑ_G, ϑ_g	ϑ_M, ϑ_m	ϑ_C, ϑ_c	ϑ_B, ϑ_b	$\vartheta_\star, \vartheta_\diamond$
Duration	τ_J, τ_j	τ_U, τ_u	τ_G, τ_g	τ_M, τ_m	τ_C, τ_c	τ_B, τ_b	$\tau_\star, \tau_\diamond$
Arbitrary singleton variable	κ_J, κ_j	κ_U, κ_u	κ_G, κ_g	κ_M, κ_m	κ_C, κ_c	κ_B, κ_b	$\kappa_\star, \kappa_\diamond$
Distribution (intensity)	Ψ^J	Ψ^U	Ψ^G	Ψ^M	Ψ^C	Ψ^B	Ψ^\star
Number of objects	N^J	...	N^G	N^M	N^C	N^B	N^\star
Haystack (space state)	W^J	W^U	W^G	W^M	W^C	W^B	W^\star
Random sample (point process)	Σ^J	Σ^U	Σ^G	Σ^M	Σ^C	Σ^B	Σ^\star
Realized sample	S^J	S^U	S^G	S^M	S^C	S^B	S^\star
Temporal rate per star	I^J	I^M	I^C	I^B	...
Abundance per star	Ξ^J	Ξ^M	Ξ^C	Ξ^B	...
Total temporal (formation) rate	\tilde{I}^J	\tilde{I}^M	\tilde{I}^C	\tilde{I}^B	\tilde{I}^\star
Arbitrary aggregate variable	K^J	...	K^G	K^M	K^C	K^B	K^\star

Note. The J-type objects are generic, and used for general formulae, as are the K and L types. Singleton variables (upper half) here are marked with the uppercase index for a type of object (e.g., B for broadcasts), if the variable is treated as a general parameter instead of describing a fixed object. A lowercase index (defaults given for each type) in a singleton variable indicates the variable refers to a realized object with fixed parameters. Other indices may be substituted, however, if they are defined for realized objects. Singleton variables may also specify a quantity window. Aggregate variables (lower half) can also be marked with an object selection, and a quantity window if different from the object window. By default, the windows are the ALL (∞) window, and the universe (U) as a host.

with a distribution of

$$\Psi_U^G(\mathbf{w}_G|\mathbf{w}_U) \equiv \frac{d\langle N_U^G \rangle}{d\mathbf{w}_G}(\mathbf{w}_G|\mathbf{w}_U) = \frac{d^3\langle N_U^G \rangle}{dN_G^\star d\mathbf{r}_G d\mathbf{w}_G}(\mathbf{w}_G|\mathbf{w}_U). \quad (62)$$

The two parameters I single out are the number of stars in the entire galaxy, N_G^\star , and its position \mathbf{r}_G , with \mathbf{w}_G' as a catch-all for any remaining parameters. All other things being equal, galaxies with more stars have more habitats in which intelligent life can evolve, increasing the mean galaxy-wide rate at which metasocieties appear. It takes time for intelligence to evolve, thus the cosmic epoch of the galaxy matters; distance also directly affects observables. There are many other factors that could affect the evolution and spread of ETIs: star formation history, metallicity distribution, stellar density, velocity dispersion, nuclear activity, and the nature of the interstellar medium all could be important (e.g., Gonzalez et al. 2001; Lineweaver et al. 2004; Dayal et al. 2015; Di Stefano & Ray 2016; Gowanlock 2016; Balbi & Tombesi 2017; Lingam et al. 2019; Lacki 2021).

It is also possible the development of new metasocieties is affected by their predecessors, and thus, there are techno-historical parameters for each galaxy as well. The science-fiction trope wherein relics of a long-dead ETI yield advanced technologies for practical interstellar travel (e.g., the “subway” system in Sagan 1985) comes to mind—later ETIs would have an easier time spreading than the first ones. Or, early ETIs might instead launch self-replicating probes that actively destroy all new ETIs (see Brin 1983, and discussion therein). Even so, we might regard all these ETIs as being part of a single metasociety, albeit one whose societies are highly clustered in time, their appearances poorly described by a Poisson process (see Section 6.3).

Observational programs only draw a sample of the cosmic population. This sample is fixed in targeted observations typical of most extragalactic radio SETI efforts to date (Shostak et al. 1996; Gray & Mooley 2017). Large-scale cosmic surveys may be viewed as drawing random samples from the Universe (Martínez & Saar 2002). Random samples of background galaxies are also present in targeted pointings at nearby objects (see Garrett & Siemion 2023). When the galaxies being studied are already known, we conditionalize on the known galaxy sample; when treating galaxies statistically, there is a galactic contribution to the sampling variance.

5.3. Stars: A Tracer of ETIs?

Planetary systems around stars are commonly regarded as the origin sites of ETIs. The properties of a sun can bear on whether a metasociety originates there.²⁹ The star needs to live long enough for intelligence to evolve, for one (Huang 1959; Carter 1983; Livio 1999), and its luminosity should be steady enough that a planet can remain habitable for geological times (Hart 1979; Kasting et al. 1993). There is a continuing debate on whether planets around M dwarfs can develop life, against the backdrop of planetary tidal locking and large stellar flares (Shields et al. 2016 and references therein); the space weather around them might discourage radio communications if a society evolves (Čirković & Vukotić 2020). Stellar parameters could also affect whether a planetary system is settled: mass and metallicity are correlated with number and types of attendant planets (e.g., Johnson et al. 2010; Mulders et al. 2015); exotic types of stars might draw ETIs interested in astroengineering (Dyson 1963; Learned et al. 2012;

²⁹ Hypothetically, ETIs could originate elsewhere; rogue planets detached from any star may even be habitable (e.g., Stevenson 1999; Abbot & Switzer 2011; Badescu 2011). On the whole, most baryonic bodies like planets should trace the stellar mass on large scales, modulated by other quantities, as discussed below.

Chennamangalam et al. 2015; Semiz & Oğur 2015; Osmanov 2016; Imara & Di Stefano 2018; Lacki 2020; Lingam & Loeb 2020), and so on. For a more precise, fine-grained treatment, we consider stellar systems instead of stars, with the distribution including quantities describing the architecture of the system: how many stars, in what orbits; the number, sizes, and locations of planets, their satellites, and other variables that could be relevant to whether there are inhabitants (e.g., Ward & Brownlee 2000).

Societies and their broadcasts may be concentrated in regions of high stellar density, without necessarily being hosted by stars. On a galactic scale, we may expect baryonic resources, stellar energy, and interstellar habitats like interstellar objects to trace the stellar mass distribution. These could be useful for ETIs in astroengineering projects. However, ETIs could migrate to the cores or outskirts of galaxies, driven by the possibility of more efficient collaboration (Kardashev 1985; Smart 2009), more optimal thermodynamic environments (Ćirković & Bradbury 2006), or large-scale gradients in habitats. Some possible attractor objects trace young stars instead of total stellar mass, like high-mass X-ray binaries (as in Vidal 2011), maybe even concentrating ETIs into a few clumps while leaving the rest of the galaxy empty. Even if they roughly follow stars, the dependence need not be linear; they might be far more prevalent in regions of higher stellar density because interstellar travel is easier, for example (Di Stefano & Ray 2016; Gajjar et al. 2021; Lacki 2021). In the absence of secure knowledge of the relevant tradeoffs, I posit that the density of societies in inhabited galaxies follows the stars on a large scale.

Although they do not fit neatly into the tree of objects, it is useful to treat stars as another class of objects with its own point process. The stars are points scattered in a *stellar haystack* W^* , described by parameter tuples w_* . Like other types of objects, windows select stars based on their position in their haystack. Galaxies host stars, but stars might also be regarded as hosted by metasocieties and societies if their planetary system is home to an ETI.³⁰ Thus, we can consider the stellar sample $\Sigma_{x,J}^*$ drawn by a selection (x, J) , consisting of $N_{x,J}^*$ stars.

The stellar properties considered here are current position r_* , birth time ϑ_* , and initial mass m_* . Lifespan τ_* is regarded as a function of m_* . The stars then have a distribution function of

$$\Psi_J^*(w_*|w_J) = \frac{d^4 \langle N_J^* \rangle}{dr_* d\vartheta_* dm_* dw_*} (w_*|w_J), \quad (63)$$

a joint stellar density, initial mass function, and star formation history. Any other parameters are absorbed into w_*' . This distribution can be used as a term in metasocietal, societal, or broadcast distribution to introduce a dependence on stellar properties.

6. The ETIs Themselves: Metasocieties and Societies

6.1. Metasocieties

Intelligences are organized into metasocieties and societies. An inhabited galaxy contains one or more metasocieties at the

present moment: $N_{t,G}^M(t) \geq 1$. A metasociety is a collection of interacting ETI societies shaped by a shared history. A society and all of its offspring are always part of the same metasociety for our purposes here. The technosignatures of a metasociety have a shared influence from a common origin. Metasocieties, unlike individual societies, can be extended over galactic scales.

The nature and evolution of metasocieties is one of the most speculative areas of SETI, with direct bearing on the Fermi paradox and expected technosignatures. A metasociety might consist of only one planet-confined society, a loose galactic network of worlds, an expanding interstellar settlement front, or a dense equilibrium state, among other possibilities (Section 6.3).

Given the diversity of hypotheses, huge uncertainties, and the possible amplification of stochastic effects by initial exponential growth, there is no one obviously correct single set of parameters to describe them. I characterize them with these parameters:

1. All metasocieties are regarded as having a single origin, arising at a single well-defined time ϑ_M , at a single location. The origin site is indicated by a parameter tuple w_M^M describing the first inhabited host star (see Section 5.3). The full distribution includes a rate of metasociety appearance per star in an empty galaxy, even allowing us to establish a functional dependence on stellar properties. However, the number of current metasocieties does not necessarily scale with $\langle N_{t,G}^* \rangle$ if they are sufficiently common and cover the entire galaxy (Section 6.3). Actual metasocieties may form as a fusion of smaller metasocieties (Kardashev 1985; Forgan 2017); the *origin* might be for the first or perhaps a dominant component, but maybe the origins of the others affect the result as well.
2. All metasocieties cover a finite volume v_M containing a finite number of astrophysical objects centered at position r_M . Not all regions with the same spatial volume are equivalent—maybe there are more societies in a volume that includes the core or disk of a galaxy than one that only includes the distant halo. An alternative is the mean number of stars covered by a metasociety, $\langle N_M^* \rangle$. In the (perhaps simplistic) scenarios I discuss, however, the metasociety is either restricted to one stellar system with $v_M \rightarrow 0$ or pervades the entire galaxy. Conceivably, multiple metasocieties might overlap, perhaps inhabiting different habitats (e.g., planetary surfaces versus deep space versus compact objects).
3. The lifespan of the metasociety, τ_M . Metasocieties may have effectively undefined lifespans, however, especially if they cover entire galaxies, continuously regenerating through internal replication.³¹
4. The mean fraction of societies within the metasociety that are communicative, f_M^C .
5. Other parameters describing its internal distribution of communicative societies, w_M^C , and metasociety-wide properties of broadcasts w_M^B . Most obviously, this can include something that governs the number of

³⁰ Formally, we could use projections of the position and time of the host (meta)society to pick out stars (Section 3.2.1). In any case, because societies are localized in single planetary systems or interstellar space, N_G^* is zero, one, or perhaps up to a few in multiple star systems, depending on whether we are considering stars per se or stellar systems.

³¹ Of course, in the very long run, all metasocieties will perish in our current understanding of cosmology, but there will be hydrogen-burning stars to host them for 10 trillion yr (Laughlin et al. 1997), much less more exotic remnants (Adams & Laughlin 1997).

communicative societies, like the rate of appearance of new societies per star. A metasociety necessarily has at least one society over its history, but the number of communicative societies may be zero if most societies are unable or unwilling to make broadcasts. For broadcasts, this could include a typical energy scale or typical rate per society. These are all lumped under a residual tuple, \mathbf{w}'_M .

Under the assumption that metasocieties are defined by these specific properties, the adopted metasocietal distribution has the form

$$\Psi_G^M(\mathbf{w}_M|\mathbf{w}_G) = \frac{d^7 \langle N_G^M \rangle}{d\vartheta_M d\tau_M d\mathbf{r}_M dv_M df_M^C d\mathbf{w}_*^M d\mathbf{w}'_M} (\mathbf{w}_M|\mathbf{w}_G). \quad (64)$$

Depending on the scenario, metasocieties may interact in complicated ways, so the metasocietal point process $\Sigma_G^M(\mathbf{w}_G)$ can be far from Poissonian (see discussion in Section 6.3).

Although most of these parameters are regarded as single numbers, the framework is sufficiently general that they can be replaced with subtuples, expanding the haystack with more dimensions. Metasocietal size is treated here as a fixed value, if only as a placeholder because the volume does not come into play in this paper. While interstellar expansion can be quick on cosmological timescales, it still takes time and we could catch it in progress (Zackrisson et al. 2015). To account for that, we could replace *size* with parameters describing the velocity of the expansion front (Jones 1981), stellar diffusion parameters (Carroll-Nellenback et al. 2019), initial low filling factors due to percolation (Landis 1998), and so on. The origin star's parameter tuple could be expanded to model the star's trajectory through the galaxy, including its initial velocity and perhaps artificial propulsion (Badescu & Cathcart 2006), if we regard its position as remaining important. The communicative fraction f_M^C could be replaced with parameters for a function describing communicativeness as a function of societal age (Sagan 1973). Many such complications can be postulated, and the distribution altered to treat them.

6.2. Societies

In this series, a society is considered a localized, independent entity with its infrastructure.³² A communicating society is one that emits broadcasts. In this view, a *society* does not just include sentient beings, but the technology they produce. Thus, a society may produce broadcasts long after its inhabitants have perished; as many have noted, the longevity of technosignatures themselves determines the number of detectable ETIs (e.g., Carrigan 2012; Balbi & Ćirković 2021; Wright et al. 2022). Conversely, a communicating society may effectively disappear when it loses interest in broadcasting, even while the society itself thrives (Sagan 1973; Smart 2009).

Classically, SETI considers single planetary systems or worlds as unitary bodies, naturally imposing a discreteness. The society level could be much finer, though, including individual interstellar vehicles and inhabited interstellar objects, or distinct organizations on a single world. The most relevant factor here is the induced clumping of broadcasts; a vast number of interstellar transmitting entities could reduce the

Poisson fluctuations of inhabited galaxies. However, there could plausibly be more levels, each imposing its discreteness effects, requiring more complex models.³³

Each society starts at time ϑ_C and survives for a finite duration τ_C . It is located at \mathbf{r}_C ; when considering very long timescales, we can instead use parameters describing its trajectory. Thus, I adopt

$$\Psi_M^C(\mathbf{w}_C|\mathbf{w}_M) = \frac{d^4 \langle N_M^C \rangle}{d\vartheta_C d\tau_C d\mathbf{r}_C d\mathbf{w}'_C} (\mathbf{w}_C|\mathbf{w}_M) \quad (65)$$

as the societal distribution for a metasociety M. Any other parameters, including those describing a society's broadcast population, are collected into \mathbf{w}'_C . As with metasocieties, individual parameters listed here can be replaced with sets of parameters (e.g., describing the society's trajectory instead of just its position). If a metasociety has more than one society, the appearance of those societies is assumed to be Poissonian. Furthermore, all the properties of the societies are assumed to be independent of one another, including their broadcast distributions.

Long-lived interstellar metasocieties are plausibly in an equilibrium state, motivating a single Ξ_M^C for each one. If ETIs tend to fill available habitats, then they can do so in much less than 1 Gyr. Sustainability arguments suggest that metasocieties necessarily have to self-regulate to avoid resource depletion (Fogg 1987; Haqq-Misra & Baum 2009). The reality may be more complicated than a simple Poissonian equilibrium process, however, with many hierarchical levels of organization that this model neglects. Perhaps metasocieties are highly chaotic, with a turbulent series of spikes, booms, plateaus, and crashes. Metasocieties that consist of several submetasocieties knit together by messages in relics could be extremely intermittent. Percolation hypotheses suggest that metasocieties may be spatially inhomogeneous as well (Landis 1998), albeit subject to stellar mixing (Wright et al. 2014) and societal turnover (Wiley 2011). Without interstellar travel, even a multitude of independent transmitting programs is *clumped* onto the rare broadcasting host worlds that presumably dominate the Poissonian fluctuations. More detailed models are necessary to treat effects like these.

6.3. Scenarios for Metasocietal Evolution

6.3.1. The Classical Scenario: Metasocieties as Societies

The default scenario in SETI theory is that societies are confined to one solar system, surviving for a continuous span of time before disappearing forever without replication (e.g., Oliver & Billingham 1971). Thus, each metasociety covers just one stellar system and contains just one society, which may or may not be communicative. Each (meta)society is regarded as appearing and behaving independently of the others. The

³² Though *society* may suggest distinct individuals cooperating toward common goals, the concept applies even if these (inter)planetary entities are solitary beings, group minds, ecospheres, or even large collections of noninteracting intelligences.

³³ What are called *civilizations* or *societies* in SETI are planetary metasocieties; all the diverse cultures, institutions, and peoples of a planet are amalgamated into one entity. This single entity could have a much longer lifespan than any society as sociologically defined. But a single world or planetary system can be far more tightly integrated than an entire galaxy, so there is motivation to treat them as one unit.

adopted metasocietal distribution function is

$$\Psi_G^M(\mathbf{w}_M|\mathbf{w}_G) = \delta(\mathbf{r}_M - \mathbf{r}_*^M) \delta(v_M) \psi[\tau_M, f_M^C, \mathbf{w}_M']_G \cdot \tilde{I}_*^M(v_M|\mathbf{w}_*^M) \Psi_G^*(\mathbf{w}_*^M|\mathbf{w}_G). \quad (66)$$

The metasociety thus has zero size ($v_M = 0$) and its position is identical to that of the star it originated from. The number of metasocieties is proportional to the stellar distribution, but it is modulated by the number rate of metasocieties appearing around each star at a given time after its birth. That term includes dependence on the stellar parameter tuple—factors like the star’s position in a galaxy, mass, and metallicity might all affect the prevalence of metasocieties. Note also that the delay-time distribution does not generally integrate to 1; it is vastly less if ETIs are rare, but can be more than 1 if technological societies evolve several times in a given planetary system (Wright 2018; Schmidt & Frank 2019).³⁴

With only one society per metasociety, there is no coordination or dependence between different societies. Whether or not a metasociety’s single society is communicative is random

$$N_M^C(\mathbf{w}_M) \sim \text{Bern}(f_M^C), \quad (67)$$

with a probability f_M^C determined by \mathbf{w}_M , the *communicative* factor in the Drake equation. The distribution of metasocieties (and societies) is entirely set by galactic properties. For a realized galaxy they are treated as a Poissonian process because of this independence (as in Glade et al. 2012):

$$N_{x,G}^M(\mathbf{w}_G) \sim \text{Pois}(\langle N_{x,G}^M \rangle) \\ N_{x,G}^C(\mathbf{w}_G) \sim \text{Pois}(\langle N_{x,G}^C \rangle) \sim \text{Pois}(\langle f_M^C \rangle_{x,G} \langle N_{x,G}^M \rangle), \quad (68)$$

where the marking and thinning of societies by their communicativeness preserves the Poisson character. Communicative societies inherit the lifespan and origin point of their parent metasociety, leading to a societal distribution

$$\Psi_M^C(\mathbf{w}_C|\mathbf{w}_M) = f_M^C \delta(\vartheta_C - \vartheta_M) \delta(\tau_C - \tau_M) \delta(\mathbf{r}_C - \mathbf{r}_M) \psi[\mathbf{w}_C']_M. \quad (69)$$

As long as $\langle \tau_M \rangle_G$ is much shorter than the galaxy’s evolutionary timescales,

$$\langle N_{t,G}^M(t) \rangle \approx \Gamma_G^M(t) \langle \tau_M \rangle_G \langle N_{t,G}^*(t) \rangle \\ \langle N_{t,G}^C(t) \rangle \approx \Gamma_G^M(t) \langle \tau_M \rangle_G \langle f_M^C \rangle_G \langle N_{t,G}^*(t) \rangle. \quad (70)$$

Thus, $\Xi_G^M \approx \Gamma_G^M \langle \tau_M \rangle_G$. The rate and abundance themselves depend on factors like star formation rate and the delay-time distribution (Section 6.4).

6.3.2. Galactic Clubs as Metasocieties

In the galactic club scenario, the first communicative society successfully contacts the succeeding ones through either remote transmissions or nonreplicating probes, establishing norms and

protocols (Bracewell 1975). The protocols become locked in, with each new communicative society conforming to the precedent set by the previous ones. The implicit coordination results in a single galaxy-spanning metasociety, even though the societies themselves remain planet bound, aside possibly from some automated probes. The classical scenario is a special kind of galactic club scenario; it simply interprets *metasociety* differently.

Much of the metasocietal distribution function is a formality. A metasociety’s size and position are simply those of the host galaxy. A galactic club is born with the first society, its origin star is the site of that society, and it endures for τ_∞ , an arbitrary timescale longer than the current age of the galaxy. The number of metasocieties in a given host galaxy is a Bernoulli random variable, simply determined by the societies $P(N_{t,G}^M(t) = 1) = P(N_{h,G}^C(t) \geq 1)$. The way the metasociety influences its children is through a single metasociety-wide f_M^C , and the distribution of any additional parameters in \mathbf{w}_M' describing shared societal and broadcast properties.³⁵

The galactic club metasociety is therefore largely an emergent phenomenon, considered here only as a mechanism to coordinate societal and broadcast properties. Societies are the fundamental entities driving the galactic club. Despite the seemingly circular dependence between the metasociety and its hosted societies, in practice, the dependence is one way for any single parameter. The births of societies are entirely independent of each other, so if we restrict the societal haystack to just \mathbf{r}_C and ϑ_C , a realized host galaxy’s population is a Poisson point process. Meanwhile, the metasociety’s f_M^C and \mathbf{w}_M' are chosen, and then used to mark each society with its other parameters. The societal haystack is then expanded into the remaining dimensions, and thinned according to which ones are communicative. Finally, the origin of the metasociety is identified with the first (communicative) society.

When we consider a host galaxy with fixed properties, and conditionalize on any metasocietal parameters governing societal/broadcast distributions, the societies are modeled with a Poisson point process:

$$N_{x,G}^C(\mathbf{w}_G|\mathbf{w}_M') \sim \text{Pois}(\langle N_{x,G}^C \rangle_{\mathbf{w}_M'}). \quad (71)$$

The societal distribution includes a dependence on the stellar distribution because they originate around stars:

$$\Psi_G^C(\mathbf{w}_C|\mathbf{w}_G, \mathbf{w}_M') = \psi[\tau_C, \mathbf{w}_C']_G \cdot \int_{\mathbf{w}_*} \delta(\mathbf{r}_C - \mathbf{r}_*) \tilde{I}_*^C(\vartheta_C|\mathbf{w}_*, \mathbf{w}_M') \Psi_G^*(\mathbf{w}_*|\mathbf{w}_G, \mathbf{w}_M') d\mathbf{w}_*. \quad (72)$$

Actual galactic metasocieties may be more complicated, even in the absence of interstellar travel. Forgan (2017) argues that instead multiple *galactic cliques* would arise, although they may in turn contact each other and fuse through a long period of inter-negotiation. There may in fact be a patchwork of different technosignature footprints in a galaxy, on intermediate spatial scales.³⁶

6.3.3. Expansive Metasocieties

Common interstellar travel allows for galaxy-spanning metasocieties. In an expansive metasociety, the metasociety

³⁴ A simple societal intensity distribution does not contain enough information to model the clustering of metasocieties that happens when they evolve multiple times on rare worlds (e.g., those with complex multicellular life). An additional level for inhabited worlds can be added to the tree to address that. Nor does Equation (66) account for what happens if subsequent metasocieties around the same star overlap; it is presumed that their lifespans are much shorter than the gap time between them.

³⁵ For example, perhaps each society broadcasts only at a frequency assigned by the metasociety according to some distribution. Or perhaps the culture of the galactic club can bias the lifespans of its member societies.

³⁶ This might also be true of expansive metasocieties.

spreads throughout a galaxy in a (cosmologically) minuscule time, treated as instantaneous here.³⁷ The technosignature population of the galaxy effectively undergoes a *phase transition* (see Ćirković & Vukotić 2008). Only one metasociety is allowed at a time in a galaxy. Like the galactic club, the size and position of the metasociety is that of the galaxy itself. A metasociety may persist once established as internal migration reestablishes societies in locations that have fallen, in which case we can set $\psi[\tau_M]_G = \delta(\tau_M - \tau_\infty)$. Or, perhaps some internal process causes them to vanish in mere millions of years or less (as in Haqq-Misra & Baum 2009; Prantzos 2020). Note this lifespan is likely far greater than that of individual societies; a galaxy-wide catastrophe is required for a galaxy-spanning metasociety to collapse.

Each metasociety arises from a single stellar system, but because only one metasociety exists at a time, the realized rate can have a complicated dependence. If a new society evolves independently in an already established metasociety, there are many possible outcomes: it could be assimilated into it without affecting its properties appreciably, it could rejuvenate a metasociety and extend its lifespan, or perhaps it could subsume the extant metasociety with its own. There could be different metasocieties coexisting, or aggressive metasocieties could inhibit the evolution of intelligence on all other planets in the galaxy, and so on. Each of these affects the metasocietal distribution in different ways. The key point is that this interaction induces a *dependent* thinning on the metasocietal point process.

A fairly simple subscenario is one where metasocieties are exclusive and inhibitory: once a metasociety appears, it persists without interference for its entire lifespan, then vanishes. No other metasociety can arise during its reign. This means $N_{t,G}^M(\vartheta_M, \mathbf{w}_G)$ is a Bernoulli variable of 1 if the galaxy has a metasociety at $t = \vartheta_M$ and 0 if it is empty. To model these metasocieties, define an auxiliary distribution describing their appearance rate in the absence of inhibition, if $N_{t,G}^M(\vartheta_M)$ just happened to be 0:

$$\begin{aligned} \Psi_G^{M=0}(\mathbf{w}_M|\mathbf{w}_G) &= \delta(\mathbf{r}_M - \mathbf{r}_G) \delta(v_M - v_G) \psi[\tau_M, f_M^C, \mathbf{w}_M]_G \\ &\cdot \check{I}_\star^{M=0}(\vartheta_M|\mathbf{w}_\star^M) \Psi_G^*(\mathbf{w}_\star^M|\mathbf{w}_G), \end{aligned} \quad (73)$$

with $\check{I}_\star^{M=0} = d\langle N_\star^M | N_{t,G}^M(\vartheta_M) = 0 \rangle_G / d\vartheta_M$. In an unpopulated galaxy, metasocieties originate around stars much like they do in the classical scenario. Without any interaction, their population is Poisson given the galaxy's properties. Hence, there is again a stellar distribution function and a rate per star. This counterfactual distribution can be marginalized to get an effective rate of appearance $\check{I}_G^{M=0} \equiv d\langle N_G^M | N_{t,G}^M(\vartheta_M) = 0 \rangle / d\vartheta_M$. If extant metasocieties have lifespans unaffected by any independent second origins in their domains, then the distribution function is

$$\Psi_G^M(\mathbf{w}_M|\mathbf{w}_G) = (1 - \langle N_{t,G}^M(\vartheta_M) \rangle) \cdot \Psi_G^{M=0}(\mathbf{w}_M|\mathbf{w}_G). \quad (74)$$

It is found by integrating Ψ_G^M with origin time ϑ_M restricted by duration τ_M ($t - \tau_M \leq \vartheta_M \leq t$; Section 3.8). Differentiating by ϑ_M turns this into a partial differential equation.

³⁷ If a galaxy is seeded with life by direct panspermia (Crick & Orgel 1973), in turn evolving ETIs, it could be viewed as hosting an expansive metasociety that takes billions of years to develop.

A few simple results are evident. When every expansive metasociety is persistent, $\tau_M \rightarrow \tau_\infty$. A galaxy then has an expansive metasociety if and only if one has ever appeared at any time during its history. The only way it lacks one is if no society develops during any interval of time. Because the first society in a realized galaxy arises according to a Poisson process, the time until one appears has an exponential distribution governed by $\Gamma_G^{M=0} \equiv \check{I}_G^{M=0} / \langle N_{t,G}^*(\vartheta_M) \rangle$. The mean number of metasocieties is

$$\langle N_{t,G}^M(t) \rangle = 1 - \exp(-\langle N_{t,G}^*(t) \rangle \Xi_G^{M=0}(t|\mathbf{w}_G)), \quad (75)$$

defined in terms of the effective abundance per star

$$\Xi_G^{M=0}(t|\mathbf{w}_G) = \frac{1}{\langle N_{t,G}^*(t) \rangle} \int_{\vartheta_G}^t \check{I}_G^{M=0}(\vartheta_M|\mathbf{w}_G) d\vartheta_M, \quad (76)$$

in turn, using the effective rate found by marginalizing $\Psi_G^{M=0}$.

But what if metasocieties are short-lived compared to the galaxy's age? If $\Gamma_G^{M=0}(\vartheta_M)$ is roughly constant given a \mathbf{w}_G , then the evolution should be described by

$$\frac{d\langle N_{t,G}^M(t) \rangle}{dt} \approx \Gamma_G^{M=0} \langle N_{t,G}^*(t) \rangle (1 - \langle N_{t,G}^M(t) \rangle) - \frac{\langle N_{t,G}^M(t) \rangle}{\langle \tau_M \rangle_G}, \quad (77)$$

with the first term on the right side describing the modified appearance rate, and the second term describing their disappearance. The natural equilibrium number of metasocieties is

$$\langle N_{t,G}^M(t) \rangle = [1 + (\Gamma_G^{M=0} \langle N_{t,G}^*(t) \rangle \langle \tau_M \rangle_G)^{-1}]^{-1}. \quad (78)$$

The convergence to 1 in Equation (78) is slower because all galaxies start with $N_{t,G}^M(t) = 0$. If $\langle \tau_M \rangle_G$ is long on cosmological scales, then there has not been time to reach the second gap between metasocieties—each galaxy has either avoided having any or is in the era of its first metasociety. In contrast, if $\langle \tau_M \rangle_G$ is short, then each galaxy has already had a random sequence of metasocieties and gaps. Equation (78) thus describes the duty cycle of expansive metasocieties once the transient initial condition has vanished. I adopt Equation (75) unless otherwise stated, but to order of magnitude the abundance of metasocieties should be the same.

Expansive metasocieties differ from galactic clubs in that the number of communicative societies saturates at a very high equilibrium level once they become established—potentially in the billions or more. We have a metasociety-level societal distribution of

$$\begin{aligned} \Psi_M^C(\mathbf{w}_M|\mathbf{w}_M) &= \Gamma_M^C(\vartheta_C|\mathbf{w}_M) \psi[\tau_C, \mathbf{w}_C']_M \\ &\cdot \frac{d\langle N_{t,M}^*(\vartheta_C) \rangle}{d\mathbf{r}_\star}(\mathbf{r}_C|\mathbf{w}_M). \end{aligned} \quad (79)$$

Furthermore, Γ_M^C is not only expected to be larger than but *independent* of $\Gamma_G^{M=0}$. The number of communicative societies in an expansive metasociety is treated separately. They form a Poisson point process, dependent on the metasociety's properties:

$$N_{x,M}^C(\mathbf{w}_M) \sim \text{Pois}(\langle N_{x,M}^C \rangle) \quad (80)$$

with $\langle N_{x,M}^C \rangle \approx \Gamma_M^C \langle \tau_C \rangle_M \langle N_{x,G}^* \rangle$ when $T_x \ll \tau_C$. This means $N_{t,G}^M(t)$ has a mixture distribution, with a $(1 - \langle N_{t,G}^M(t) \rangle)$ probability of being 0 and an $\langle N_{t,G}^M(t) \rangle$ probability of being

$N_{t,M}^C(t)$, yielding

$$P(N_{t,G}^C(t) \geq 1) = \langle N_{t,G}^M(t) \rangle [1 - \exp(-\langle N_{t,M}^C(t) \rangle_G)]. \quad (81)$$

The independence results in interesting effects: in particular, when $\langle N_{t,G}^M(t) \rangle \ll 1$, the probability of detecting a broadcast can vary quadratically with the number of stars in a galaxy (Section 12).

The qualitative characteristics of galactic clubs and persistent expansive metasocieties converge when $\langle N_G^M \rangle$ approaches 1. In both cases, metasociety-wide traits can be treated as galactic traits and societies form a Poisson point process in a realized galaxy. A galactic club can even be regarded as an expansive metasociety that spreads through telecommunications instead of settlement, albeit with much lower population densities in the absence of replication. A single metasociety approximation, with $N_G^M = 1$, serves to work for both scenarios when ETIs are not very rare.

6.4. The Drake Equation in the Formalism

The Drake equation applies to the classical scenario when the galaxy is in a steady state (Drake 1965; Glade et al. 2012). New stars are formed at a constant rate, and as they age over billions of years, some fraction of those host single ETI societies of limited lifespan. To implement it simplistically, if an ETI evolves, it must appear exactly \bar{t} after the star's birth: the delay-time distribution is

$$\check{I}_*^M = f_*^M \delta((\vartheta_M - \vartheta_*) - \bar{t}). \quad (82)$$

Stars need to maintain a stable luminosity long enough for intelligence to evolve on Earth-like planets, excluding massive stars. However, let us imagine for simplicity that all stars with $m_* \leq m_*^H$ are equally likely to host an ETI. This probability f_*^M includes all terms in the Drake equation relating to the number of planets, habitability, and the evolution of life and intelligence. Then we can apply the above to the classical metasociety distribution (Equation (73)).

We want to calculate the number of communicative societies existing at any one time in the galaxy. First, let us calculate the galactic-level societal distribution, using Equation (15). We convolve $\Psi_M^C(\mathbf{w}_C|\mathbf{w}_M)$ with $\Psi_G^M(\mathbf{w}_M|\mathbf{w}_G)$. Directly plugging these into that equation yields a formidable integral over all stellar and metasocietal parameters. Actually, most of the variables are effectively nuisance parameters, and the expression is simplified by the many delta functions and the separable integrals. In the end, we are left with

$$\begin{aligned} \Psi_G^C(\mathbf{w}_C|\mathbf{w}_G) = & \langle f_*^M \rangle_G \langle f_M^C \rangle_G \cdot \psi[\mathbf{r}_*^M]_G(\mathbf{r}_C) \cdot \psi[\tau_M]_G(\tau_C) \\ & \cdot \psi[\mathbf{w}_C]_G \cdot \check{I}_G^*(\vartheta_M - \bar{t}) \cdot F[m_*^M](m_*^H). \end{aligned} \quad (83)$$

In other words, the communicative societies trace the stars in space, each existing for the lifespan of its metasociety, and they appear at a rate proportional to the star formation rate \bar{t} ago, including only stars that last long enough for ETIs to evolve.

The expected number of currently active communicative societies $\langle N_{t,G}^C(t) \rangle$ is the integral of Ψ_G^C with $t - \tau_C \leq \vartheta_C \leq t$ (Section 3.8). Now let us suppose that the star formation rate of the galaxy has been constant, that (meta)societies are always much younger than the galaxy, and that $F[m_*^M](m_*^H) \approx 1$ because most stars are low mass and likely to survive for the

required billions of years. Then we finally get

$$\langle N_{t,G}^C(t) \rangle \approx \check{I}_G^* \langle f_*^M \rangle_G \langle f_M^C \rangle_G \langle \tau_M \rangle_G. \quad (84)$$

This is the Drake equation rewritten in the formalism.

Including different delay-time distributions, galactic habitability, dependence on stellar mass, and other effects is fairly straightforward. We simply change the $d\langle N_*^M \rangle_G/d\vartheta_M^*$ distribution to include dependencies on these parameters of the host star. A spread in delay times is necessary when treating quiescent galaxies where the star formation rate was quenched about 10 billion yr ago.

7. Broadcasts: General Considerations

A broadcast is an artificial release of energy at a discrete site through some specified time and frequency range (see Tables 6–8 for notations). Not every technosignature is a broadcast: unpowered artifacts in the solar system (Freitas & Valdes 1985; Rose & Wright 2004; Davies & Wagner 2013) and anomalous atmospheric compositions resulting from industrial pollution (Whitmire & Wright 1980; Lin et al. 2014) are not broadcasts. On the other hand, not every broadcast need be an attempt at communication—they include directed energy transmission for power beaming (Inoue & Yokoo 2011; Benford & Benford 2016), propulsion (Lingam & Loeb 2017), or remote sensing (Scheffer 2014). Broadly speaking, waste heat from megastructures (Dyson 1960) and exhaust radiation from vehicles (Harris 1986) are broadcasts too.

The modulation of starlight by astronomical-scale megastructures (Arnold 2005; Chennamangalam et al. 2015; Wright et al. 2016; Zackrisson et al. 2018; Lacki 2019; Suazo et al. 2022) is an interesting case—here, the technosignature is the blocking of an energy release that normally would happen. These might be viewed as negative energy broadcasts. Searches for galactic-scale obscuration of starlight (Annis 1999; Zackrisson et al. 2015) use a collective bound as in Paper II, much as waste heat searches look for the positive energy release from the reprocessing of this missing starlight.

7.1. Basic Parameters to Describe Broadcasts

The basic considerations for whether a broadcast is detectable are where it is, how bright it becomes, when it happens, and how it behaves. The first question is answered by the broadcast position \mathbf{r}_B .

Fundamentally, the brightness is controlled by the total energy released ϵ_B and its distribution into the transmitter's sky over solid angle ω_B (refer to Table 8 for emission-related variables). These can be considered separately, but I use the total effective isotropic energy $\hat{\epsilon}_B \equiv 4\pi\epsilon_B/d\omega_B$ of the broadcast to describe the brightness, evaluated toward the direction of the observer. In the absence of beaming, $\hat{\epsilon}_B = \epsilon_B$. When broadcasts are beamed, the $\hat{\epsilon}_B$ distribution has a large peak near zero for f -axis broadcasts. Yet even on-axis beamed broadcasts may have a wide range of $\hat{\epsilon}_B$ because of different beaming angles or simply different ϵ_B . Previous SETI works have essentially considered two basic classes of $\hat{\epsilon}_B$ distributions—the monoenergetic δ -distribution with a characteristic $\hat{\epsilon}_B$ and power-law distributions (e.g., Drake et al. 1973; Gulkis 1985; Dreher 2004). Both narrow and broad distributions seem plausible at this point, with the former favored by possible engineering constraints on deliberate transmissions and the

Table 7
Other Variable Notation for Objects

Notation	Explanation
Galaxies and Stars	
m_*	Mass of individual star
$M_{t,G}^*$	Total stellar mass of galaxy at one epoch
Metasocieties and Societies	
f_M^C	Fraction of societies in a metasociety M that are communicative (broadcasting)
w_*^M	Parameter tuple for the star from which the metasociety first emerged
$\dot{\gamma}_*^M$	Mean rate of metasocietal origin events around a star
$\dot{\gamma}_*^C$	Mean rate of societal origin events around a star
Broadcasts	
ϖ_B	Quantity describing polarization of broadcast; in general, it is a vector or matrix
$\zeta_{p B}(p)$	Fraction of broadcast emission in polarization p
$\zeta_{x B}$	Sum of $\zeta_{p B}$ over all polarizations in Π_x
ω_B	Solid angle that broadcast is emitted into
β_B	Total bandwidth of broadcast in the source frame, frequency span over the entire duration
$\beta_{t B}$	Instantaneous bandwidth of broadcast in the source frame
$\beta_{x:t B}$	Bandwidth covered by broadcast over the duration of x
ν_B	Central frequency of broadcast in the source frame
δ_B	Drift rate of broadcast in the source frame
Λ_J^B	Mean temporal rate of broadcasts in host J per unit frequency per star
Z_J^B	Mean number of broadcasts in host J per unit frequency per star
$\dot{\Lambda}_J^B$	Mean total temporal rate of broadcasts in host J per unit frequency
\dot{Z}_J^B	Mean total number of broadcasts in host J per unit frequency
\oplus	Superscript indicating observer-frame quantity

latter motivated by the possible diversity of ETI capabilities and goals. Each broadcast's emission may also have distinct degrees and states of polarization. The polarization properties are encapsulated in ϖ_B , which should be understood as a Stokes vector or similar representation.

Surveys are sensitive to particular kinds of time–frequency properties—spectral lines or pulses, for example. Although the time–frequency behavior may be endlessly complex, I model it with five basic parameters. Broadcasts are limited in time, with a source-frame duration τ_B , and in frequency, with a constant source-frame instantaneous bandwidth $\beta_{t|B}$. Broadcasts also start at times ϑ_B and are centered on source-frame frequencies ν_B . Finally, a drift rate δ_B describes the skewness of the broadcast in the time–frequency space. The broadcast is entirely contained in the time range $[\vartheta_B, \vartheta_B + \tau_B]$ and frequency range $[\nu_B - \beta_B/2, \nu_B + \beta_B/2]$, where the total bandwidth β_B is derived from $\beta_{t|B}$ and δ_B ($\beta_B = \beta_{t|B} + \tau_B|\delta_B|$ when they are constant).

These quantities define a haystack in which broadcasts are scattered. Most of them match the quantities defining haystacks in Tarter (2007) and Wright et al. (2018). A separate modulation parameter is missing here, not being directly relevant for energy or photon detection, but bandwidth and drift rate both can be considered kinds of modulation. The main differences are in the treatment of time, reduced to a single quantity in previous works. Wright et al. (2018) interpret this as a repetition period, while the duty cycle and the longevity of

the transmitter are ignored. Including periodicity would necessitate adding at least one more dimension to the ones here (see discussion in Section 4.1). Bursts repeating at random intervals (Kipping & Gray 2022), however, require no new dimensions—the mean rate is simply considered a property of the hosting society.

7.2. Selection of Broadcasts

When is a broadcast sampled? Primary considerations are whether a broadcast touches the window in physical spacetime (Section 3.8) and frequency, with a cross section set by duration and bandwidth. A broadcast has an effective bandwidth $\beta_{x:t|B}$ during window x , a quantity accounting for a broadcast's spectral evolution. In some cases, the broadcast's spectrum is unchanging (like the box model; Section 9), in which case we can simply use $\beta_{x:t|B} = \beta_{t|B}$. For the purposes of selection, I adopt

$$\beta_{x:t|B} = \text{diam} \left\{ \nu \left| \frac{d^2 \tilde{\epsilon}_{x|B}}{dt d\nu}(t, \nu | w_B) > 0 \right. \right\}, \quad (85)$$

where $\text{diam } A = \max[|a_1 - a_2|, a_1, a_2 \in A]$.³⁸

I regard the selection of a broadcast by window x as a binary decision based on its time and frequency properties:

$$P_x^B(w_B) = \mathbb{I}[-\tau_B \leq \vartheta_B - \Theta_x \leq \tau_B] \cdot \mathbb{I}[|\nu_B - \Upsilon_x| \leq (1/2)(\beta_{x:t|B} + B_x)]. \quad (86)$$

The presumption is that the broadcast's ancestors also are chosen by x , as required (Section 3.2.3) because if the broadcast is active at some time and place, so are all its hosts—so x must not impose any filters on characteristics specific to society, metasociety, or galaxy. Polarization characteristics are not considered because of the likely crosstalk between the broadcast's polarization and the observed polarizations, at least for linear polarization.

7.3. Energy and Photons Emitted by Broadcasts

A broadcast is detected through its emission (Table 8). This can be measured in various forms, including energy and photons. Each broadcast has an (effective isotropic) energy output per polarization per unit time and frequency

$$\dot{\ell}_{\nu,p;B}(t, \nu, p) \equiv \frac{d^2 \tilde{\epsilon}_{p;B}}{dt d\nu}(t, \nu, p) = \zeta_{p|B}(p) \dot{\ell}_{\nu;B}(t, \nu). \quad (87)$$

I assume that the degree and state of polarization are independent of time and frequency. Thus, I separate $\tilde{\epsilon}_{\nu,p;B}$ into the fraction of energy emitted in polarization p , $\zeta_{p|B}(p)$, and the spectral luminosity, $\dot{\ell}_{\nu;B}$.³⁹ The same basic ideas can be applied to broadcasts in other messengers, like neutrinos and gravity waves.

Of course, we do not collect all of the energy emitted by a broadcast, only that which falls on our instruments within a time and frequency window in a set of observed polarizations.

³⁸ $\beta_{x:t|B}$ and $\beta_{x|B}$ are distinct—the latter accounts for the frequency, sky field, and so on as well as restricting the bandwidth. A broadcast at a frequency never covered by x has $\beta_{x|B} = 0$ and a drifting signal that just barely touches x has $\beta_{x|B} \ll B_x$, but neither situation affects $\beta_{x:t|B}$. In all cases, $\beta_B \geq \beta_{x:t|B} \geq \beta_{x|B}$.
³⁹ For example, a broadcast with linear polarization fraction $\varpi_{\text{lin};B}$ has $\zeta_{p|B}(p_{\text{lin}}) = (1/2)(1 - \varpi_{\text{lin};B}) + \varpi_{\text{lin};B} \cos^2 \phi_{p_{\text{lin}};B}$, where $\phi_{p_{\text{lin}};B}$ is the relative angle between broadcast's polarization angle and p_{lin} .

Table 8
Notation for Emission Quantities

Quantity	Arbitrary	Energy	Power	Photons	Note
Emission from broadcast	\mathcal{Q}_B	ϵ_B	ℓ_B	q_B	a,b
Effective isotropic emission from broadcast	$\hat{\mathcal{Q}}_B$	$\hat{\epsilon}_B$	$\hat{\ell}_B$	\hat{q}_B	b
... per unit frequency	$\hat{\mathcal{Q}}_{\nu;B}$	$\hat{\epsilon}_{\nu;B}$	$\hat{\ell}_{\nu;B}$	$\hat{q}_{\nu;B}$	c
... per polarization p	$\hat{\mathcal{Q}}_{p;B}$	$\hat{\epsilon}_{p;B}$	$\hat{\ell}_{p;B}$	$\hat{q}_{p;B}$	c
... per unit frequency per polarization p	$\hat{\mathcal{Q}}_{\nu,p;B}$	$\hat{\epsilon}_{\nu,p;B}$	$\hat{\ell}_{\nu,p;B}$	$\hat{q}_{\nu,p;B}$	c
... occurring during window x	$\hat{\mathcal{Q}}_{x B}$	$\hat{\epsilon}_{x B}$	$\hat{\ell}_{x B}$	$\hat{q}_{x B}$	c
Effective isotropic aggregate emission from broadcast population selected by (x, J)	$\hat{\mathcal{R}}_{x,J}^B$	$\hat{E}_{x,J}^B$	$\hat{L}_{x,J}^B$	$\hat{Q}_{x,J}^B$	b
Effective isotropic background emission in window x	$\hat{\mathcal{Q}}_{x n}$	$\hat{\epsilon}_{x n}$	$\hat{\ell}_{x n}$	$\hat{q}_{x n}$	b
Effective isotropic total emission in window x from object J	$\hat{\mathcal{R}}_{x,J}$	$\hat{E}_{x,J}$	$\hat{L}_{x,J}$	$\hat{Q}_{x,J}$	b
Emission distance	$d_{g;B}$	$d_{e;B}$	$d_{\ell;B}$	$d_{q;B}$	
Dilution factor for broadcast	$y_{g;B}$	$y_{e;B}$	$y_{\ell;B}$	$y_{q;B}$	
Transmittance factor within window x	$\chi_{g;x B}$	$\chi_{e;x B}$	$(\chi_{\ell;x B})$	$\chi_{q;x B}$	
Fluence (flux) from a single broadcast in window x	$u_{x B}$	$h_{x B}$	$f_{x B}$	$g_{x B}$	d
Aggregate fluence (flux) from a population of broadcasts selected by (x, J)	$\mathcal{U}_{x,J}^B$	$\mathcal{H}_{x,J}^B$	$\mathcal{F}_{x,J}^B$	$\mathcal{G}_{x,J}^B$	d
Background fluence (flux) in window x	$u_{x n}$	$h_{x n}$	$f_{x n}$	$g_{x n}$	d
Total fluence (flux) including background in window x from object J	$\mathcal{U}_{x,J}$	$\mathcal{H}_{x,J}$	$\mathcal{F}_{x,J}$	$\mathcal{G}_{x,J}$	d
Instrument-measured quantity for single broadcast in window x	$m_{x B}$	$e_{x B}$	$l_{x B}$	$q_{x B}$	
Aggregate measured quantity from a population of broadcasts selected by (x, J)	$\mathcal{M}_{x,J}^B$	$\mathcal{E}_{x,J}^B$	$\mathcal{L}_{x,J}^B$	$\mathcal{Q}_{x,J}^B$	
Background measured quantity in window x	$m_{x n}$	$e_{x n}$	$l_{x n}$	$q_{x n}$	
Total instrument-measured quantity in window x	\mathcal{M}_x	\mathcal{E}_x	\mathcal{L}_x	\mathcal{Q}_x	

Notes.

^a When no quantity window is given, the ALL window is assumed to apply; so $\hat{\epsilon}_B$ is the total effective isotropic energy released over the entire lifespan, frequency range, and all polarizations from broadcast B.

^b The variable listed under power is the effective isotropic radiated power (EIRP).

^c Any emission variable, aggregate, or singleton, can be substituted with these modifiers to the same effect. Thus, $\hat{L}_{p;x,J}^B$ is the aggregate single-polarization EIRP from the broadcasts in $\Sigma_{x,J}^B$; $\hat{f}_{\nu;B}$ is the energy flux per unit frequency of a single random broadcast B, unfiltered by any window.

^d The variable listed under power is energy flux; all others are types of fluence.

Out of a total effective isotropic emission $\hat{\mathcal{Q}}_B$, the broadcast releases only a limited amount $\hat{\mathcal{Q}}_{x|B}$ within these constraints, zero if the broadcast is not in the sample. The amount of energy released within the time, frequency, and polarization constraints for x is

$$\hat{\epsilon}_{x|B} = \int_{\Upsilon_x - B_x/2}^{\Upsilon_x + B_x/2} \int_{\Theta_x}^{\Theta_x + T_x} \sum_{p \in \Pi_x} \hat{\ell}_{\nu,p;B}(t, \nu, p) dt d\nu. \quad (88)$$

When the polarization properties are independent of time and frequency, the sum over polarizations is replaced by a single factor $\zeta_{x|B}$, such that

$$\zeta_{x|B} = \sum_{p \in \Pi_x} \zeta_{p|B}(p|\varpi_B). \quad (89)$$

Note that the frequency and time variables are all in the *source frame*.⁴⁰

When dealing with photon detectors, the relevant quantity is the number of photons emitted within the observational window

$$\hat{q}_{x|B} \equiv \int_{\Upsilon_x - B_x/2}^{\Upsilon_x + B_x/2} \int_{\Theta_x}^{\Theta_x + T_x} \sum_{p \in \Pi_x} \frac{\hat{\ell}_{\nu,p;B}(t, \nu, p)}{h\nu} dt d\nu, \quad (90)$$

where h is Planck's constant.

⁴⁰ An object at redshift z is observed at Earth to have $\Upsilon_x = \Upsilon_x^\oplus(1+z)$, $B_x = B_x^\oplus(1+z)$, and $T_x = T_x^\oplus/(1+z)$.

Even though the time domain and frequency domain amplitudes are related by the Fourier transform, $\hat{\ell}_{\nu;B}$ can have extremely complex dependence on time and frequency.⁴¹ This motivates the later use of the box and chord models as simplifications.

7.4. Distances, Fluences, and Dilutions

Observable quantities depend not just on the intrinsic properties of the broadcast, but where it is—its distance (d_B) and position in the sky (θ_B). The difference in redshift between broadcasts in the same galaxy is insignificant, so I set them equal ($z_B = z_G$). Variation in broadcast distances is important for the Milky Way, which after all is another galaxy with some population of broadcasts (even if it is empty aside from our own). For extragalactic systems that are not Galactic satellites, however, the size of the target galaxy is negligible compared to its distance, and we can approximate the distance as that to the galaxy itself ($d_B = d_G$).

Aside from the cuts imposed by the selections, only the emission incident on our detector can be measured. The fluence $u_{x|B}$ of a broadcast is the amount of emission per unit area

⁴¹ Formally, β_B and τ_B cannot both be finite; applying a rectangular window to one domain convolves the other by a sinc function that is nonzero over an infinite range. Many observations integrate over $T_0 B_0 \gg 1$, resulting in negligible leakage. These include all practical observations in the optical or higher energies, and filterbank data products in radio SETI which also employ polyphase filterbank techniques to reduce sidebands (Price 2021).

integrated over all combinations of time, frequency, and polarization covered by window \mathbf{x} . Specific types of fluence include the energy fluence $h_{\mathbf{x}|\mathbf{B}}$ where the emission is quantified as energy and the photon fluence $g_{\mathbf{x}|\mathbf{B}}$, which counts the expected number of photons (Table 8).

The distance determines the flux and fluence of a broadcast with a given amount of emission. The energy fluence can be found using the *energy distance* $d_e = \sqrt{1+z}d_M$, which accounts for redshift, correcting the transverse angular distance d_M (see Hogg 1999 for how to calculate d_M). Likewise, the photon fluence uses a *photon distance* d_q equal to d_M .

The observed emission is also partly blocked by opacity from dust and gas on the sightline, plus the Earth's atmosphere, a range-limiting issue in optical SETI (Howard et al. 2004). Only a fraction $\chi_{\varrho;\mathbf{x}|\mathbf{B}}$ of ϱ emission remains after absorption and scattering, the transmittance factor, which can vary with both position and frequency. The frequency dependence means that windows at different frequencies have different transmittances. At the usual radio frequencies observed in SETI, however, $\chi_{\varrho;\mathbf{x}|\mathbf{B}}$ is very near 1.

It is convenient to express the inverse-square law and redshift effects as a *dilution factor*, $y_{\varrho;\mathbf{B}} = 1/(4\pi d_{\varrho;\mathbf{B}}^2)$. For emission of type ϱ ,

$$u_{\mathbf{x}|\mathbf{B}} = \frac{\dot{\varrho}_{\mathbf{x}|\mathbf{B}} \chi_{\varrho;\mathbf{x}|\mathbf{B}}}{4\pi d_{\varrho;\mathbf{B}}^2} = \dot{\varrho}_{\mathbf{x}|\mathbf{B}} \chi_{\varrho;\mathbf{x}|\mathbf{B}} y_{\varrho;\mathbf{B}}. \quad (91)$$

In principle, the $\dot{\varrho}_{\mathbf{x}|\mathbf{B}}$ and $y_{\varrho;\mathbf{B}}$ are not independent—although the intrinsic properties of the broadcast are modeled as independent, the cuts depend on \mathbf{T}_x and \mathbf{B}_x , which are affected by redshift. Even the Doppler shifts of broadcasts should have a slight effect on $\dot{\varrho}_{\mathbf{x}|\mathbf{B}}$. Transmittance, when it is not near 1, is likely to be much lower in regions of high obscuration, and thus, is certainly not independent of $y_{\varrho;\mathbf{B}}$. However, for practical purposes, with the emission and dilution independent,

$$\langle u_{\mathbf{x}|\mathbf{B}}^n \rangle_{\mathbf{x},\mathbf{J}} = \langle \dot{\varrho}_{\mathbf{x}|\mathbf{B}}^n \rangle_{\mathbf{x},\mathbf{J}} \langle \chi_{\varrho;\mathbf{x}|\mathbf{B}}^n y_{\varrho;\mathbf{B}}^n \rangle_{\mathbf{J}} \quad (92)$$

for any arbitrary exponent n .

8. Populations of Broadcasts

8.1. The Broadcast Distribution Function

In this series, the broadcast distribution function takes the form

$$\Psi_{\mathbf{J}}^{\mathbf{B}}(\mathbf{w}_{\mathbf{B}}|\mathbf{w}_{\mathbf{J}}) = \frac{d^8 \langle N_{\mathbf{J}}^{\mathbf{B}} \rangle}{d\hat{e}_{\mathbf{B}} d\varpi_{\mathbf{B}} d\tau_{\mathbf{B}} d\vartheta_{\mathbf{B}} d\beta_{\mathbf{t}|\mathbf{B}} dv_{\mathbf{B}} d\delta_{\mathbf{B}} d\mathbf{r}_{\mathbf{B}}}(\mathbf{w}_{\mathbf{B}}|\mathbf{w}_{\mathbf{J}}), \quad (93)$$

using the variables identified in Section 7.1. The broadcast population is a Cox point process, and the broadcast population of a realized society in particular is a Poisson point process. The instantaneous rate of broadcasts per frequency per star is especially useful:

$$\Lambda_{\mathbf{J}}^{\mathbf{B}}(t, \nu) \equiv \frac{1}{\langle N_{\mathbf{t},\mathbf{J}}^{\mathbf{B}}(t) \rangle} \frac{d^2 \langle N_{\mathbf{J}}^{\mathbf{B}} \rangle}{d\vartheta_{\mathbf{B}} d\nu_{\mathbf{B}}}(\vartheta_{\mathbf{B}} = t, \nu_{\mathbf{B}} = \nu), \quad (94)$$

as is the frequency abundance per star,

$$Z_{\mathbf{J}}^{\mathbf{B}}(t, \nu) = \frac{1}{\langle N_{\mathbf{t},\mathbf{J}}^{\mathbf{B}}(t) \rangle} \frac{d \langle N_{\mathbf{t},\mathbf{J}}^{\mathbf{B}}(t) \rangle}{d\nu_{\mathbf{B}}}(\nu_{\mathbf{B}} = \nu). \quad (95)$$

The stellar rate $\Gamma_{\mathbf{J}}^{\mathbf{B}}$ and the stellar abundance $\Xi_{\mathbf{J}}^{\mathbf{B}}$ are additionally marginalized over $\nu_{\mathbf{B}}$, as defined as in Section 3.2.4. It is more convenient to use the total abundances and rates, the derivatives without dependence on $\langle N_{\mathbf{t},\mathbf{J}}^{\mathbf{B}}(\nu) \rangle$, when considering individual societies around single star systems: $\check{\Lambda}_{\mathbf{C}}^{\mathbf{B}} \equiv d^2 \langle N_{\mathbf{C}}^{\mathbf{B}} \rangle / d\vartheta_{\mathbf{B}} dv_{\mathbf{B}}$, $\check{Z}_{\mathbf{C}}^{\mathbf{B}} \equiv d \langle N_{\mathbf{t},\mathbf{C}}^{\mathbf{B}}(t) \rangle / dv_{\mathbf{B}}$, $\check{\Gamma}_{\mathbf{C}}^{\mathbf{B}} \equiv d \langle N_{\nu,\mathbf{C}}^{\mathbf{B}}(\nu) \rangle / d\vartheta_{\mathbf{B}}$, and $\check{\Xi}_{\mathbf{C}}^{\mathbf{B}} \equiv \langle N_{\mathbf{t} \cap \nu, \mathbf{C}}^{\mathbf{B}}(t, \nu) \rangle$.

Specific examples of the broadcast distribution are given in Sections 9 and 10.

8.2. Interchangeability

When I compute properties of aggregate observables in this paper, I apply the results for compound Poisson random variables. However, to do this, the variables being summed need to be identically distributed. Basically, each society is assumed to have the same broadcast distribution as the others, an average distribution. Likewise, each metasociety has the same societal and broadcast distributions as any other in the galaxy. The *interchangeability* assumption is that all distributions of an object type are the same, except for translations in spacetime (because objects are located in different places and start at different times). The broadcast distribution in every society (regardless of metasociety) in galaxy \mathbf{G} has the form

$$\Psi_{\mathbf{C}}^{\mathbf{B}}(\mathbf{w}_{\mathbf{B}}|\mathbf{w}_{\mathbf{C}}) = f_{\mathbf{G}}^{\mathbf{B}}(\hat{e}_{\mathbf{B}}, \varpi_{\mathbf{B}}, \tau_{\mathbf{B}}, \beta_{\mathbf{t}|\mathbf{B}}, \nu_{\mathbf{B}}, \delta_{\mathbf{B}}) \cdot \delta(\mathbf{r}_{\mathbf{B}} - \mathbf{r}_{\mathbf{C}}) \cdot \mathbb{I}[0 \leq \vartheta_{\mathbf{B}} - \vartheta_{\mathbf{C}} \leq \tau_{\mathbf{C}}], \quad (96)$$

and every metasociety has a societal distribution

$$\Psi_{\mathbf{M}}^{\mathbf{C}}(\mathbf{w}_{\mathbf{C}}|\mathbf{w}_{\mathbf{M}}) = f_{\mathbf{G}}^{\mathbf{C}}(\tau_{\mathbf{C}}) \cdot \mathbb{I}[0 \leq \vartheta_{\mathbf{C}} - \vartheta_{\mathbf{M}} \leq \tau_{\mathbf{C}}], \quad (97)$$

with fixed functions $f_{\mathbf{G}}^{\mathbf{B}}$ and $f_{\mathbf{G}}^{\mathbf{C}}$.

Societies are selected at different points of their life cycles, but for the compound Poisson distribution to be applicable, the societal broadcast distribution must be time stationary. If, for example, societies lose interest in communicating as they age, then young societies have a higher $\Lambda_{\mathbf{C}}^{\mathbf{B}}$ than older ones. Interchangeability does not allow this. Thus, there can be no further dependence on $\vartheta_{\mathbf{B}} - \vartheta_{\mathbf{C}}$ in Equation (96), as long as the broadcast is within bounds. Likewise, the societal distribution is stationary. More generally, there cannot be any edge effects wherein societies or metasocieties spend different amounts of time in the selection window; thus, all lifespans must either all be much longer or much shorter than the window's duration.

Real societies and metasocieties may be incredibly diverse, each with its own unique array of technosignatures. Societies and broadcasts are represented by Cox processes (Section 2.5). Because these are generally nonergodic, a single society or metasociety is a very poor representative of the ensemble properties of the underlying distribution—we only see one realization of the distribution, which itself is a random variable. If half of societies broadcast in radio and the rest in gamma rays, and there is only a single society that happens to be radio broadcasting, we will not be able to see the true diversity in broadcasts. The observed sample is thus subjected to stochasticity when there are only a few societies. A large sample of societies can sample the entire gamut of broadcasts. In that sense, all the broadcasts can be pooled together into a galactic broadcast distribution, although the effects of the clumping remain underestimated by the compound Poisson distribution.

There are expressions for the mean and variances when dealing with diverse populations, but they are very complicated

Table 9
Equations for Aggregate Variables Under Standard Assumptions

Variable	Galactic Club	Expansive Metasociety
Societal Variables		
$P(N_{x,C}^B \geq 1)$	$1 - \exp(-\langle N_{x,C}^B \rangle)$	$1 - \exp(-\langle N_{x,C}^B \rangle)$
$\mathbb{V}[N_{x,C}^B]$	$\langle N_{x,C}^B \rangle$	$\langle N_{x,C}^B \rangle$
$\langle K_{x,C}^B \rangle$	$\langle N_{x,C}^B \rangle \langle \kappa_{x B} \rangle_C$	$\langle N_{x,C}^B \rangle \langle \kappa_{x B} \rangle_C$
$\mathbb{V}[K_{x,C}^B]$	$\langle N_{x,C}^B \rangle \langle \kappa_{x B}^2 \rangle_C$	$\langle N_{x,C}^B \rangle \langle \kappa_{x B}^2 \rangle_C$
Metasocietal Variables		
$P(N_{x,M}^C \geq 1)$...	$1 - \exp(-\langle N_{x,M}^C \rangle)$
$\mathbb{V}[N_{x,M}^C]$...	$\langle N_{x,M}^C \rangle$
$\langle K_{x,M}^C \rangle$...	$\langle N_{x,M}^C \rangle \langle \kappa_{x M}^C \rangle$
$\mathbb{V}[K_{x,M}^C]$...	$\langle N_{x,M}^C \rangle \langle \kappa_{x M}^2 \rangle_C$
$P(N_{x,M}^B \geq 1)$...	$1 - \exp[-\langle N_{x,M}^C \rangle (1 - e^{-\langle N_{x,C}^B \rangle})]$
$\langle N_{x,M}^B \rangle$...	$\langle N_{x,M}^C \rangle \langle N_{x,C}^B \rangle$
$\mathbb{V}[N_{x,M}^B]$...	$\langle N_{x,M}^C \rangle [1 + \langle N_{x,C}^B \rangle]$
$\langle K_{x,M}^B \rangle$...	$\langle N_{x,M}^C \rangle \langle \kappa_{x B} \rangle_C$
$\mathbb{V}[K_{x,M}^B]$...	$\langle N_{x,M}^C \rangle [\langle \kappa_{x B}^2 \rangle_C + \langle N_{x,C}^B \rangle \langle \kappa_{x B} \rangle_C^2]$
Galactic Variables		
$P(N_{x,G}^C \geq 1)$	$1 - \exp(-\langle N_{x,G}^C \rangle)$	$\langle N_{x,G}^M \rangle [1 - \exp(-\langle N_{x,G}^C \rangle)]$
$\langle N_{x,G}^C \rangle$	$\langle N_{x,G}^C \rangle$	$\langle N_{x,G}^M \rangle \langle N_{x,M}^C \rangle$
$\mathbb{V}[N_{x,G}^C]$	$\langle N_{x,G}^C \rangle$	$\langle N_{x,G}^C \rangle [1 + \langle N_{x,M}^C \rangle (1 - \langle N_{x,M}^C \rangle)]$
$\langle K_{x,G}^C \rangle$	$\langle N_{x,G}^C \rangle \langle \kappa_{x C} \rangle$	$\langle N_{x,G}^M \rangle \langle N_{x,M}^C \rangle \langle \kappa_{x C} \rangle$
$\mathbb{V}[K_{x,G}^C]$	$\langle N_{x,G}^C \rangle \langle \kappa_{x C}^2 \rangle$	$\langle N_{x,G}^C \rangle [\langle \kappa_{x C}^2 \rangle + \langle N_{x,M}^C \rangle (1 - \langle N_{x,M}^C \rangle) \langle \kappa_{x C} \rangle^2]$
$P(N_{x,G}^B \geq 1)$	$1 - \exp[-\langle N_{x,G}^C \rangle (1 - e^{-\langle N_{x,C}^B \rangle})]$	$\langle N_{x,G}^M \rangle [1 - \exp(-\langle N_{x,M}^C \rangle (1 - e^{-\langle N_{x,C}^B \rangle}))]$
$\langle N_{x,G}^B \rangle$	$\langle N_{x,G}^C \rangle \langle N_{x,C}^B \rangle$	$\langle N_{x,G}^M \rangle \langle N_{x,M}^C \rangle \langle N_{x,C}^B \rangle$
$\mathbb{V}[N_{x,G}^B]$	$\langle N_{x,G}^C \rangle [1 + \langle N_{x,C}^B \rangle]$	$\langle N_{x,G}^M \rangle [1 + (\langle N_{x,C}^B \rangle + (1 - \langle N_{x,M}^C \rangle) \langle N_{x,M}^C \rangle)]$
$\langle K_{x,G}^B \rangle$	$\langle N_{x,G}^C \rangle \langle \kappa_{x B} \rangle_C$	$\langle N_{x,G}^M \rangle \langle \kappa_{x B} \rangle_C$
$\mathbb{V}[K_{x,G}^B]$	$\langle N_{x,G}^C \rangle [\langle \kappa_{x B}^2 \rangle_C + \langle N_{x,C}^B \rangle \langle \kappa_{x B} \rangle_C^2]$	$\langle N_{x,G}^M \rangle [\langle \kappa_{x B}^2 \rangle_C + \langle \kappa_{x B} \rangle_C^2 (\langle N_{x,C}^B \rangle + (1 - \langle N_{x,M}^C \rangle) \langle N_{x,M}^C \rangle)]$

Note. These expressions are calculated under the assumptions of independence and interchangeability. Note that $\langle N_{x,K}^J \rangle_L \equiv \langle N_{x,K}^J \rangle_{x,L}$ and $\langle \kappa_{x|J} \rangle_K \equiv \langle \kappa_{x|J} \rangle_{x,K}$ by convention (Section 3.4.2).

because of the nesting selection-relative averages and variances (Equation (45)). It is more insightful to work directly with the distributions themselves, when they are available (Section 8.6).

Table 9 lists equations for some important quantities characterizing broadcast populations, given the assumptions of independence and interchangeability.

8.3. The Number of Broadcasts

The number of broadcasts intercepted by \mathbf{x} from a society $C \in \Sigma_{x,J}$ is $N_{x,C}^B(\mathbf{w}_C)$, a Poisson random variable given \mathbf{w}_C . Its mean is given by the number of broadcasts that touch the \mathbf{x} window:

$$\langle N_{x,C}^B \rangle = \int \dots \int_{\Upsilon_x - (\beta_{x:|B} + \mathbf{B}_x)/2}^{\Upsilon_x + (\beta_{x:|B} + \mathbf{B}_x)/2} \int_{\Theta_x - \tau_B}^{\Theta_x + \tau_x} \Psi_C^B(\mathbf{w}_B | \mathbf{w}_C) d\vartheta_B dv_B d\epsilon_B d\varpi_B d\tau_B d\beta_{|B} d\mathbf{r}_B d\delta_B. \quad (98)$$

If societies are long-lived compared to τ_B , and if Λ_C^B varies slowly enough in time and frequency compared to common values of τ_B and $\beta_{x:|B}$,

$$\langle N_{x,C}^B \rangle \approx \check{\Lambda}_C^B(\Theta_x, \Upsilon_x) \cdot (\mathbf{B}_x + \langle \beta_{x:|B} \rangle_C) (\tau_x + \langle \tau_B \rangle_C). \quad (99)$$

The number of broadcasts is a compound random sum, with means and variances given in the table. As long as $\tau_B \ll \tau_C, \tau_M$ and Λ_G^B is slowly varying,

$$\langle N_{x,G}^B \rangle \approx \Lambda_G^B(\Theta_x, \Upsilon_x) \cdot \langle N_{x,G}^* \rangle (\mathbf{B}_x + \langle \beta_{x:|B} \rangle_G) \cdot (\tau_x + \langle \tau_B \rangle_G). \quad (100)$$

It is scaled by the mean rate that broadcasts occur per unit frequency per star, which is

$$\Lambda_G^B = \begin{cases} \Xi_G^C \langle \check{\Lambda}_C^B \rangle_G & \text{(galactic club)} \\ \langle N_{x,G}^M \rangle \langle \Xi_M^C \rangle_G \langle \check{\Lambda}_C^B \rangle_G & \text{(expansive interstellar).} \end{cases} \quad (101)$$

This abundance is a key quantity in SETI. The clumping into societies and metasocieties adds additional sampling variance, more so in the expansive interstellar scenario where $N_{x,G}^M$ is independent of $N_{x,M}^C$. When fewer than one broadcast is typically selected per society, the probability that one broadcast

is intercepted is

$$P(N_{x,G}^B \geq 1) \approx \begin{cases} 1 - \exp(-\langle N_{x,G}^B \rangle) & \text{(galactic club)} \\ \langle N_{x,G}^M \rangle [1 - \exp(-\langle N_{x,M}^B \rangle)] & \text{(expansive interstellar).} \end{cases} \quad (102)$$

In the opposite limit of many selected broadcasts per society, this probability is simply the probability that one communicative society is selected. The variance in the number of broadcasts per metasociety is then mainly from the Poissonian clumping into societies, approaching $\langle N_{x,M}^C \rangle \langle N_{x,C}^B \rangle^2$ (full expressions are given in Table 9).

8.4. Aggregate Emission and Fluence

The total single-polarization spectral luminosity from the broadcasts in a sample $\Sigma_{x,J}^B$ is

$$\dot{L}_{\nu,p;x,J}^B(t, \nu, p) = \sum_{w_B \in \Sigma_{x,J}^B} \dot{\ell}_{\nu,p;B}(t, \nu, p). \quad (103)$$

Of course, $\Sigma_{x,J}^B$ itself is random, even when J itself is specified. The usual conditions for broadcast selection by the window x (Equation (86)) define the bounds of integration to be used in Equation (41). Under the interchangeability-independence assumptions,

$$\begin{aligned} \langle \dot{L}_{\nu,p;x,J}^B(t, \nu, p) \rangle &\approx \Lambda_J^B(t, \nu) \cdot \langle N_{x,J}^* \rangle \langle \dot{\ell}_{\nu,p;B}(t, \nu, p) \rangle_{x,J} \\ &\cdot (B_x + \langle \beta_{x;t|B} \rangle) (T_x + \langle \tau_B \rangle_J). \end{aligned} \quad (104)$$

Strictly speaking, this is not necessarily a compound Poisson variable because $N_{x,G}^M(w_G)$ is Bernoulli in the expansive metasociety scenario. It reaches the compound Poisson limit when $\langle N_{x,G}^M \rangle = 1$ and the galaxy as a whole can be identified with the metasociety (Table 9).

The aggregate fluence is a similar compound sum under these assumptions. A fluence $\mathcal{U}_{x,J}^B$ of any type from the sample is found by summing fluences $u_{x|B}$ of the same type from individual broadcasts in a sample. Examples include the total energy fluence $\mathcal{H}_{x,J}^B$ and total photon fluence $\mathcal{G}_{x,J}^B$ of a sample, composed of the energy fluences $h_{x|B}$ and photon fluences $g_{x|B}$ of individual broadcasts. We have

$$\begin{aligned} \langle \mathcal{U}_{x,J}^B \rangle &\approx \Lambda_J^B(\Theta_x, \Upsilon_x) \cdot \langle N_{x,J}^* \rangle (B_x + \langle \beta_{x;t|B} \rangle) \\ &\cdot (T_x + \langle \tau_B \rangle_J) \langle u_{x|B} \rangle_J. \end{aligned} \quad (105)$$

The clumping of broadcasts into societies can have a big effect on the variance of $\mathcal{U}_{x,J}^B$ (Table 9). When a window typically samples $\ll 1$ broadcast per society or the variance in $u_{x|B}$ is sufficiently large, the variance in an inhabited galaxy G with $N_{x,G}^M = 1$ approaches what we would expect if $N_{x,G}^B$ were Poissonian. In the opposite limit, $\mathbb{V}[\mathcal{U}_{x,G}^B] \approx \langle N_{x,G}^B \rangle \langle N_{x,C}^B \rangle \langle u_{x|B/C} \rangle^2 = \langle N_{x,G}^C \rangle \langle N_{x,C}^B \rangle^2$. The societies become a population of standard candles, and the fluctuations are just shot noise in their number. This limit requires many broadcasts per society just within a window, all becoming confused, which can be a tall order (Paper II).

8.4.1. What Do We Expect the Aggregate Emission to Look Like?

Although we usually envision broadcasts to resemble isolated spikes in \dot{L}_{ν} , on coarse enough time–frequency scales, this roughness blurs out. The mean spectrum, as shown by Equation (104), is proportional to Λ_G^B itself, as long as it varies

slowly enough and a mean $\dot{\ell}_{\nu,p;B}$ exists. If there is no magic frequency (unlike Drake & Sagan 1973; Blair & Zadnik 1993) or magic time (unlike in Pace & Walker 1975; Makovetskii 1977; Corbet 2003; Nishino & Seto 2018) that ETIs in the target galaxy prefer to broadcast at, Λ_G^B is most likely a smooth function, perhaps a power law.

Supposing that the $\dot{\ell}_B$ distribution has a mean, over large bandwidths and long durations the aggregate emission converges to a constant luminosity source with a smooth spectrum. In other words, it will resemble diffuse nonthermal emission commonly seen in objects like active galactic nuclei. Like any population of discrete sources (e.g., Tonry & Schneider 1988), there are Poissonian fluctuations in the emission, far greater than those from the immense number of particles contributing to natural diffuse emission. With enough broadcasts being added together, the fluctuations also become Gaussian by the central limit theorem. This motivates the use of the total emission of galaxies to set constraints on broadcasts, as in Paper II.

If the $\dot{\ell}_B$ distribution has no mean, or if broadcasts are so rare that confusion does not occur for any feasible sample, the aggregate emission remains spiky, with strongly non-Gaussian fluctuations that could be picked out by conventional search strategies. Nonetheless, there still could be a background of fainter broadcasts for which aggregate constraints still apply.

Even if there is a magic frequency, the broadcasts from an extended metasociety may be smeared out in received frequency because of velocity differences between the different transmitter sites. Then the aggregate emission will appear much like natural line emission, although possibly at a frequency corresponding to no natural transition. The luminosity of candidate lines in the galaxy's emission sets constraints on such broadcast populations.

8.5. The Diffuse Approximation

Because societies are discrete, they clump broadcasts. However, including the societal level in the tree greatly complicates the analysis, compounding the Poissonian character of the number of broadcasts in collections of societies. The *diffuse approximation* ignores the discreteness of societies, instead imagining the transmitters being spread diffusely across the galaxy. The variance in $N_{x,G}^B$ is thus underestimated. How large this correction is depends on the nature of the discretization, but the diffuse approximation gives good results when there are many societies, few of which emit a detectable broadcast ($\langle N_{x,C}^B \rangle_J \ll 1$). This is appropriate when using very fine observations (e.g., narrow channels for narrowband lines) except in the most extreme cases—far into the confusion regime for heavily populated galaxies.

Practically speaking, the diffuse approximation treats the broadcasts as the children of metasocieties (expansive metasociety scenario) or galaxies (classical and Galactic club scenario). Thus, the diffuse approximation works directly with the metasociety's or galaxy's broadcast distribution instead of building those up from societies' broadcast distributions. Just as the societal distribution posits that societies trace stars (Sections 5.3 and 6.3), in the diffuse approximation, that dependence is shifted down to the broadcast distribution. In the end, the broadcast distribution for the effective parent J in the diffuse approximation is found by substituting

$$\check{\Lambda}_C^B \delta(\mathbf{r}_B - \mathbf{r}_C) \rightarrow \Lambda_J^B \frac{d\langle N_{t,J}^* \rangle}{d\mathbf{r}_*}(\mathbf{r}_B); \quad (106)$$

the substitution can also apply to the broadcast abundance per frequency variable.

8.6. Aggregate Luminosities in Diverse, Non-interchangeable Hosts: An Example

Diversity among host objects increases the sampling variance of aggregate quantities, and the variance increases as it manifests in higher-level ancestors. This section presents a simple example: How would the cosmic-relative variance in an aggregate metasocietal luminosity \bar{L}_M^B change if the luminosities were shared among different types of hosts? To keep the example focused, the number distributions $N_C^B(w_C) \sim \text{Pois}(\langle N_C^B \rangle)$ and $N_M^C(w_M) \sim \text{Pois}(\langle N_M^C \rangle)$ are strictly the same among different societies and metasocieties. In case 1, every metasociety and society is interchangeable, with

$$\psi[\bar{L}_B]_C(\ell) = \exp(-\ell/\bar{\ell})/\bar{\ell}. \quad (107)$$

In case 2, all the broadcasts in an individual society have the same luminosity, but this characteristic luminosity is exponentially distributed between societies:

$$\psi[\bar{L}_B]_C(\ell) = \delta(\ell - \bar{\ell}_C) \text{ and } \psi[\bar{L}_C]_M(\ell) = \exp(-\ell/\bar{\ell})/\bar{\ell}. \quad (108)$$

In case 3, each metasociety has decreed all broadcasts that have the same luminosity, but different metasocieties decide on different luminosities, with an exponential distribution:

$$\psi[\bar{L}_B]_M(\ell) = \delta(\ell - \bar{\ell}_M) \text{ and } \psi[\bar{L}_M]_U(\ell) = \exp(-\ell/\bar{\ell})/\bar{\ell}. \quad (109)$$

These can be compared with case 0, in which all broadcasts have the same luminosity: $\psi[\bar{L}_B]_C(\ell) = \delta(\ell - \bar{\ell})$. In case 0, broadcasts, societies, and metasocieties are all interchangeable.

There are slightly different approaches to calculating the variance in each case, but it involves building up from broadcast-level averages to the cosmic level. All the broadcasts in a society are drawn from the same luminosity distribution—a degenerate one in cases 0, 2, and 3, and an exponential one in case 1. Thus, the aggregate luminosity of broadcasts in a realized society is a compound Poisson variable. In cases 0, 1, and 3, the societies are interchangeable, so we can then apply a compound Poisson distribution again. However, in case 2, the societies have different luminosity distributions depending on their $\bar{\ell}_C$. To advance to the metasocietal-level variables, we have to integrate over $\psi[\bar{L}_C]_M$, which is a sort of marginalized societal haystack (Equation (10)). Finally, in cases 0–2, all metasocieties are interchangeable, and so $\langle \bar{L}_M^B \rangle = \langle \bar{L}_M^B \rangle_U$ and $\mathbb{V}[\bar{L}_M^B] = \mathbb{V}[\bar{L}_M^B]_U$. However, in case 3, we need to find $\mathbb{V}[\bar{L}_M^B]_U = \langle (\bar{L}_M^B)^2 \rangle_U - \langle \bar{L}_M^B \rangle_U^2$ by averaging over $\psi[\bar{L}_C]_U$ (Equation (43)), effectively a metasocietal haystack.

In all four cases, the mean aggregate luminosity is the same, $\langle \bar{L}_M^B \rangle_U = \langle N_M^B \rangle \bar{\ell}$, because the mean number of broadcasts and the mean broadcast luminosity are the same. The variance,

however, increases from one case to the next:

$$\mathbb{V}[\bar{L}_M^B]_U = \begin{cases} 2\bar{\ell}^2 \langle N_M^B \rangle \left(\frac{1}{2} + \frac{1}{2} \langle N_C^B \rangle \right) & \text{(Case 0)} \\ 2\bar{\ell}^2 \langle N_M^B \rangle \left(1 + \frac{1}{2} \langle N_C^B \rangle \right) & \text{(Case 1)} \\ 2\bar{\ell}^2 \langle N_M^B \rangle (1 + \langle N_C^B \rangle) & \text{(Case 2)} \\ 2\bar{\ell}^2 \langle N_M^B \rangle \left(1 + \langle N_C^B \rangle + \frac{1}{2} \langle N_M^B \rangle \right) & \text{(Case 3)}. \end{cases} \quad (110)$$

In case 0, the variance is totally due to compound Poissonian fluctuations in the number of broadcasts, supplemented by intrinsic variance in the luminosity in case 1. All the broadcasts in a society share the same luminosity in cases 2 and 3, amplifying the fluctuations; in case 3, as the shared luminosity $\bar{\ell}_M$ varies, the entire aggregate luminosity does as well.

This example only scratches the surface of the subject, but the methods in Sections 2 and 3 can be applied to these problems even when the compound Poisson distribution does not apply.

9. The Box Model

In the box model, broadcasts are simple contiguous *boxes* in time–frequency space. Each box covers the frequency range $|\nu - \nu_B| \leq \bar{\beta}/2$ and time range $0 \leq t - \vartheta_B \leq \tau$. Because these boxes are not skewed, the drift rate δ_B is 0. Furthermore, all the boxes are assumed to have identical bandwidths $\bar{\beta}$ and durations τ (Figure 7). The distribution is assumed to be stationary, in that ϑ_B and ν_B have uniform distributions over any observable range, with

$$\Psi_C^B = \check{A}_C^B \cdot \psi[\bar{\epsilon}_B, \varpi_B]_C \cdot \delta(\tau_B - \bar{\tau}) \delta(\beta_{\text{tB}} - \bar{\beta}) \delta(\delta_B) \delta(\mathbf{r}_B - \mathbf{r}_C). \quad (111)$$

In the box model, $\beta_{x:t|B} = \bar{\beta}$ (Equations (85) and (98)).

The selection itself is also a contiguous box in time–frequency space, spanning the frequency range $|\nu - \Upsilon_x| \leq B_x/2$ and time range $0 \leq t - \Theta_x \leq T_x$. Sample properties in the box model depend on the number and amount of overlap between broadcast boxes and the sample box in time–frequency space (Figure 7).

Consider a sample $\Sigma_{x,J}^B$, consisting of broadcasts within a field of $N_{x,J}^*$ stars with a time and frequency response described by a box. The number of broadcasts expected from a host J is proportional to the rate per star per unit frequency, number of stars, and the temporal and frequency cross sections resulting from the combined width of the window and the broadcasts:

$$\langle N_{x,J}^B \rangle = A_J^B \langle N_{x,J}^* \rangle (T_x + \bar{\tau}) (B_x + \bar{\beta}). \quad (112)$$

Observables in the box model tie to an overlap quantity $o_{x|B}$, the fraction of the broadcast's time–frequency *area* within the sample box. Appendix B provides calculations relating to this quantity.

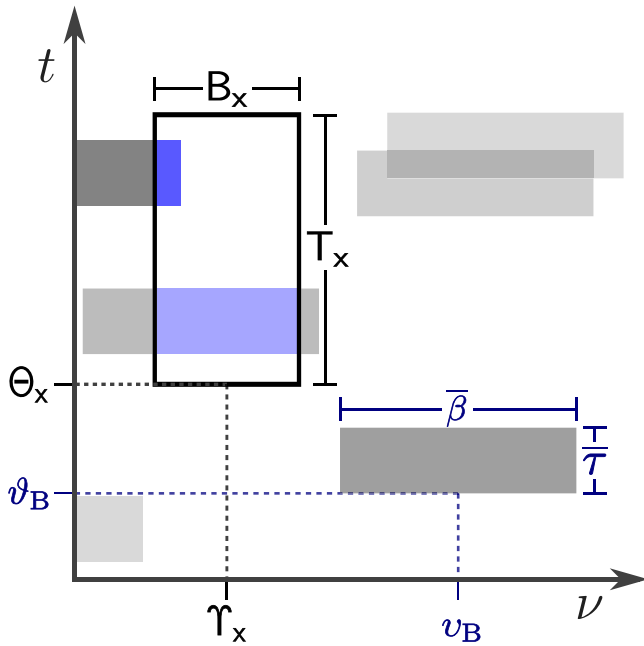


Figure 7. Sketch of a spectrogram in the box model. All broadcasts have equal duration τ and bandwidth β , and $\ell_{\nu,B}$ is constant within each box although the total ϵ_B may vary (filled boxes of different shades). The selection also is a window of duration T_x and bandwidth observation B_x at all times and frequencies within those ranges. Only emission within the selection (blue shading) is observed.

As uniform boxes in time–frequency space, broadcasts in the box model have steady luminosities and flat spectra:

$$\ell_{\nu,p;B}(t, \nu, p) = \begin{cases} \zeta_{p|B}(p) \frac{\epsilon_B}{\beta\tau} & \text{if } 0 \leq t - t_B \leq \tau \\ & \text{and } |\nu - \nu_B| \leq \beta/2 \\ 0 & \text{otherwise.} \end{cases} \quad (113)$$

Although the broadcasts may share the same degree and type of polarization, the polarization state is random. The mean luminosity spectrum of a sample is

$$\langle \ell_{\nu,p;B}^B(t, \nu, p) \rangle = \frac{1}{2} \langle N_{x,J}^* \rangle A_J^B \langle \epsilon_B \rangle_{x,J}, \quad (114)$$

with a mean energy fluence of

$$\langle \mathcal{H}_{x,J}^B \rangle = \frac{|\Pi_x|}{2} A_J^B \langle N_{x,J}^* \rangle \langle \epsilon_B \rangle_{x,J} \langle \gamma_{\epsilon;B} \chi_{\epsilon;|B} \rangle_{x,J} T_x B_x. \quad (115)$$

Many observational quantities simplify when we assume that $\tau \approx T_x$ and $\beta \approx B_x$. The box model has four natural limits:

1. *Lines*—Long-lasting, narrowband broadcasts with $\tau \gg T_x$ but $\beta \ll B_x$. These include the classic carrier wave beacons of radio SETI (Drake 1961; see also Enriquez et al. 2017 and references therein) and artificial laser lines in the optical (Schwartz & Townes 1961). Instead of ϵ_B and A_J^B , it is most natural to consider the isotropic luminosity $\ell_B \equiv \epsilon_B/\tau$. As the lifetime is unknown and largely irrelevant, the observable rate is $Z_J^B = A_J^B \langle \tau_B \rangle_J = A_J^B \tau$.
2. *Pulses*—Short-lasting, wideband broadcasts with $\tau \ll T_x$ and $\beta \gg B_x$. Several SETI surveys have sought pulses (e.g., Shvartsman et al. 1993; Howard et al. 2004; Siemion et al.

2010; Maire et al. 2019). Instead of ϵ_B , it is most natural to consider the isotropic energy spectrum $\ell_{\nu;B} \equiv \epsilon_B/\beta$. The

observable rate is $\Gamma_J^B = A_J^B \langle \beta_B \rangle_J = A_J^B \beta$.

3. *Blips*—Short-lasting, narrowband broadcasts with $\tau \ll T_x$ and $\beta \ll B_x$. Blips can form the basis of wideband communication systems (Messerschmitt 2015), or result from transient broadcasts like our radars or narrow-beam beacons on rotating worlds (Gray & Ellingsen 2002).
4. *Hisses*—Long-lasting, wideband broadcasts with $\tau \gg T_x$ and $\beta \gg B_x$, essentially noise. Thermal waste heat (Dyson 1960) is effectively a hiss, as is exhaust radiation (Harris 1986), and a typical random waveform is a hiss similar to white noise. Long-lasting continuum sources emit hisses. Instead of ϵ_B , it is most natural to consider the isotropic luminosity spectrum $\ell_{\nu;B} \equiv \epsilon_B/(\tau\beta)$. The observable abundance is $\Xi_J^B = A_J^B \langle \tau_B \rangle_J \langle \beta_B \rangle_J = A_J^B \tau \beta$.

Table 10 summarizes key quantities for each of these regimes.

Often there are multiple time and frequency scales involved in a survey. A broadcast might be longer than a few seconds long exposure but multiple pointings of the same sky region could occur over years. The approximations for the four quadrants apply when the inequalities hold for all relevant time and frequency scales.

10. The Chord Model for Frequency-drifting Lines

The broadcasts in the box model have no skew. Although a satisfactory approximation when the time or frequency resolution is coarse, both line and pulse searches use fine sampling windows sensitive to frequency drifts—lines drift because of the changing Doppler shifts of an accelerating source (Sheikh et al. 2019), while pulses drift because of dispersion induced by the interstellar and intergalactic media (Siemion et al. 2010). In fact, lines without any drift or dispersion are generally attributable to anthropogenic radio frequency interference, since the rotation of Earth will add a location-dependent frequency drift (Enriquez et al. 2017; Sheikh et al. 2021). Individual broadcasts can be dedrifted or dedispersed by shifting the spectrum by different delays in different channels to maximize signal-to-noise (see Paper II). Nonetheless, a population of broadcasts form a background with a range of drift rates that cannot be dedrifted.

An archetypal drifting line never ends in the middle of an observational sample window because $T_x \ll \tau_B$. To avoid edge effects, these broadcasts can be considered as thin bands with an instantaneous bandwidth $\beta_{t|B}$ and a duration at fixed frequency $\tau_{\nu|B}$ (Figure 8). The ratio of these quantities defines the slope or drift rate

$$|\delta_B| \equiv \frac{\beta_{t|B}}{\tau_{\nu|B}}. \quad (116)$$

The drift rate is a signed quantity that is positive if the broadcast drifts to high frequencies at later times and negative if it drifts to low frequencies. In the chord model, δ_B does not vary with time—broadcasts have linear drift in frequency. As in the box model, the selection window is a contiguous box in time–frequency space.

The chord model simplifies further by taking these bands to have negligible $\beta_{t|B}$ and $\tau_{\nu|B}$, even as δ_B itself is fixed. Because the band has infinitesimal bandwidth, the broadcast

Table 10
Quantities in the Extreme Regimes of the Box Model

Quantity	Line	Pulse	Blip	Hiss
τ	$\gg T_x$	$\ll T_x$	$\ll T_x$	$\gg T_x$
β	$\ll B_x$	$\gg B_x$	$\ll B_x$	$\gg B_x$
Natural energy variable	ℓ_B	$\tilde{\epsilon}_{\nu,B}$	$\tilde{\epsilon}_B$	$\tilde{\ell}_{\nu,B}$
Natural rate variables	$Z_J^B = \Lambda_J^B \tau$	$\Gamma_J^B = \Lambda_J^B \beta$	Λ_J^B	$\Xi_J^B = \Lambda_J^B \beta \tau$
$\alpha_{x B}$	T_x/τ	B_x/β	1	$T_x B_x/(\tau \beta)$
$\langle N_{x,J}^B \rangle$	$Z_J^B \langle N_{x,J}^* \rangle B_x$	$\Gamma_J^B \langle N_{x,J}^* \rangle T_x$	$\Lambda_J^B \langle N_{x,J}^* \rangle B_x T_x$	$\Xi_J^B \langle N_{x,J}^* \rangle$
$\langle \tilde{\epsilon}_{x B} \rangle_J$	$(\Pi_x /2) \langle \tilde{\epsilon}_B \rangle_J T_x$	$(\Pi_x /2) \langle \tilde{\epsilon}_{\nu,B} \rangle_J B_x$	$(\Pi_x /2) \langle \tilde{\epsilon}_B \rangle_J$	$(\Pi_x /2) \langle \tilde{\ell}_{\nu,B} \rangle_J B_x T_x$
$\langle \tilde{\epsilon}_{x B}^2 \rangle_J$	$\langle \zeta_{x B}^2 \rangle_J \langle \tilde{\epsilon}_B^2 \rangle_J T_x^2$	$\langle \zeta_{x B}^2 \rangle_J \langle \tilde{\epsilon}_{\nu,B}^2 \rangle_J B_x^2$	$\langle \zeta_{x B}^2 \rangle_J \langle \tilde{\epsilon}_B^2 \rangle_J$	$\langle \zeta_{x B}^2 \rangle_J \langle \tilde{\ell}_{\nu,B}^2 \rangle_J B_x^2 T_x^2$
$\langle \tilde{L}_{\nu,x,J}^B \rangle$	$\frac{ \Pi_x }{2} Z_J^B \langle N_{x,J}^* \rangle \langle \tilde{\epsilon}_B \rangle_J$	$\frac{ \Pi_x }{2} \Gamma_J^B \langle N_{x,J}^* \rangle \langle \tilde{\epsilon}_{\nu,B} \rangle_J$	$\frac{ \Pi_x }{2} \Lambda_J^B \langle N_{x,J}^* \rangle \langle \tilde{\epsilon}_B \rangle_J$	$\frac{ \Pi_x }{2} \Xi_J^B \langle N_{x,J}^* \rangle \langle \tilde{\ell}_{\nu,B} \rangle_J$
$\langle h_{x B} \rangle_J$ (distant host)	$\frac{ \Pi_x }{2} \frac{\langle \tilde{\epsilon}_B \rangle_J T_x \langle \chi_{x B} \rangle_J}{4\pi(1+z)d_{M,J}^2}$	$\frac{ \Pi_x }{2} \frac{\langle \tilde{\epsilon}_{\nu,B} \rangle_J B_x \langle \chi_{x B} \rangle_J}{4\pi(1+z)d_{M,J}^2}$	$\frac{ \Pi_x }{2} \frac{\langle \tilde{\epsilon}_B \rangle_J \langle \chi_{x B} \rangle_J}{4\pi(1+z)d_{M,J}^2}$	$\frac{ \Pi_x }{2} \frac{\langle \tilde{\ell}_{\nu,B} \rangle_J B_x T_x \langle \chi_{x B} \rangle_J}{4\pi(1+z)d_{M,J}^2}$

Note. Quantities in the source frame: broadcast rates and abundances (Z_J^B , Γ_J^B , Λ_J^B , Ξ_J^B), window definition quantities (B_x , T_x), broadcast durations and bandwidths (β , τ), broadcast emission (ℓ_B , $\tilde{\epsilon}_{\nu,B}$, $\tilde{\epsilon}_B$, $\tilde{\ell}_{\nu,B}$, $\tilde{\epsilon}_{x|B}$, $\tilde{L}_{\nu,x,J}^B$). Quantities in the observer frame: fluence ($h_{x|B}$). When considering a fixed sample of stars, the known number of stars $N_{x,J}^*$ can be substituted for $\langle N_{x,J}^* \rangle$.

is either entirely within the sample or not at any given time (gray-blue line in Figure 8). A broadcast is part of the sample if the band ever crosses the box, appearing as a chord in time–frequency diagrams. The duration of the chord’s time in the sample box is $\tau_{x|B}$, and it sets the amount of emission received. The effective bandwidth used in calculating the expected number of broadcasts in the sample is $\beta_{x,t|B} = |\delta_B| \tau_B$ (Equations (98) and (85)). The chord approximation breaks down if $\tau_{\nu|B} \gtrsim \min(T_x, B_x/|\delta_B|)$.

I adopt a distribution for drifting broadcasts similar to that for lines in the box model, with a frequency abundance of \tilde{Z}_C^B per society:

$$\Psi_C^B = \tilde{Z}_C^B \delta(\tau_B - \tau_\infty) \delta(\beta_{t|B}) \delta(\delta_B) \delta(\mathbf{r}_B - \mathbf{r}_C) \delta(\vartheta_B) \cdot \psi[\tilde{\epsilon}_B, \varpi_B, \delta_B]_C. \quad (117)$$

The origin time of each broadcast is set to $t=0$, while the duration is set to an arbitrarily long τ_∞ ; the distribution is then assumed to extend to all negative and positive frequencies with the understanding that emission at frequencies and times outside any window being used is ignored. This allows for a direct relation between $\tilde{\epsilon}_B$ and the more relevant effective isotropic luminosity $\tilde{\ell}_B = \tilde{\epsilon}_B/\tau_\infty$, with $\psi[\tilde{\ell}_B]_C = \tau_\infty \psi[\tilde{\epsilon}_B]_C$. The mean number of broadcasts within $\Sigma_{x,J}^B$ is

$$\langle N_{x,J}^B \rangle = Z_J^B \langle N_{x,J}^* \rangle (B_x + T_x |\delta_B|_J), \quad (118)$$

with an average time spent within x of

$$\langle \tau_{x|B} \rangle_J = \frac{T_x}{1 + T_x |\delta_B|_J / B_x}. \quad (119)$$

The drift rate distribution is not evenly sampled. Broadcasts with high $|\delta_B|$ are more likely to cut through the box simply because they cover more frequency, although they spend less time on average within the observational window. Calculations of statistical quantities need to take this into account (see Appendix C, Equation (C9)). Although the biasing can be extreme for individual observations in single channels, modern radio SETI surveys cover hundreds of megahertz or more and thus should sample drift rates more or less fairly.

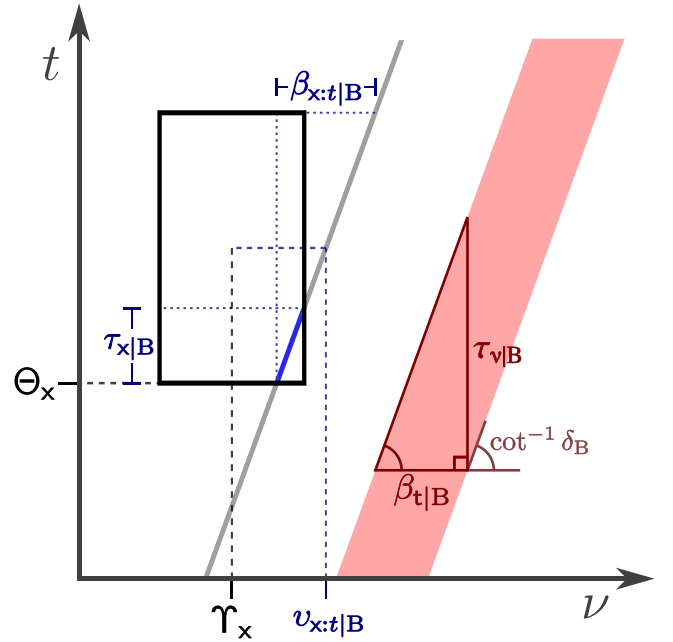


Figure 8. Sketch of a spectrogram in the chord model. Drifting signals could resemble a band (red) on the diagram. In the chord approximation, the band is assumed to have negligible bandwidth and looks like a skewed line (gray). A nonzero fluence is intercepted by the window only if the line cuts across the window box in a chord (highlighted in blue).

A drift rate distribution is necessary to calculate the higher moments. In some cases, all δ_B may have the same value, if all broadcasts are coming from transmitters at a single location within a narrow frequency range. A galactic population of line transmitters may contain a panoply of sites with different accelerations with a potentially vast δ_B range (Sheikh et al. 2019). In this series, I adopt a uniform drift rate distribution:

$$\psi[\delta_B]_J = \mathbb{I}[|\delta_B - \Delta_J^B| \leq \bar{\delta}_J] / (2\bar{\delta}_J), \quad (120)$$

where $\Delta_J^B \equiv \langle \delta_B \rangle_J$ is the center of the distribution, which may be nonzero because of Doppler effects imposed by the Earth’s rotation and revolution. A canonical value for the drift rate scale is $\bar{\delta}_J = 1 \text{ Hz s}^{-1}(\nu/\text{GHz})$ (Oliver & Billingham 1971),

though there could be a tail extending to much higher δ_B (Sheikh et al. 2019). Note this is the unbiased drift rate distribution.

The energy emission in this model is treated as being defined by a steady luminosity:

$$\dot{L}_{\nu,p;B}(t, \nu, p) = \zeta_{p|B}(p) \dot{L}_B \delta(\nu - \nu_{tB}(t)), \quad (121)$$

where $\nu_{tB}(t)$ is the frequency of the chord at time t . Taking the limit of broadcasts with small $\beta_{t|B}$ allows us to calculate

$$\langle \dot{L}_{\nu;x,J}^B(t, \nu) \rangle = \frac{|\Pi_x|}{2} \langle N_{x,J}^* \rangle Z_J^B \langle \dot{L}_B \rangle_{x,J}. \quad (122)$$

This is a smooth function although any realized $\dot{L}_{\nu;x,J}^B$ itself is very spiky. Windows integrate over these spikes with their nonzero bandwidth, yielding a mean fluence of

$$\langle \mathcal{S}_{x,J}^B \rangle = \frac{|\Pi_x|}{2} Z_J^B \langle N_{x,J}^* \rangle \langle \dot{L}_B \rangle_{x,J} \langle y_{\epsilon;B} \chi_{\epsilon;x|B} \rangle_{x,J} T_x B_x. \quad (123)$$

Further details and formulae are given in Appendix C.

11. Measurements and Noise

The final step from populations to observables is the measurement itself. A full discussion of measurements and instrumental effects is deferred (see Paper II), but it is worth noting that measurements are random variables that depend on the sample. They generally depend on a detector response \mathcal{I} that varies with position, but this distribution is predictable for each point on the haystack, and the observables are tractable with the point process framework. An archetypal observable is an aggregate variable summing a quantity $m_{x|B}$ (with $\langle m_{x|B} \rangle \propto u_{x|B}$) for every broadcast in a sample mixed with some kind of background. Measurements can also include derived quantities that describe the statistics of integrated observables. A simple example is the derived signal-to-noise ratio, which is discussed in detail in Paper II. Others include cross-correlation statistics to detect large populations of faint signals (Drake 1965). Another major type is counts of objects fulfilling a particular criterion.

Actual measurements are subject to noise, microscopic fluctuations in the instrument, background radiation, or the broadcast radiation itself. The noise increases the variance in observables beyond the sampling variance, making individual broadcasts harder to detect. If we are doing a measurement on a broadcast sample Σ_x^B , the variance in that measurement is (Equation (2))

$$\mathbb{V}[m_{x|B}] = \langle \mathbb{V}[m_{x|B} | \Sigma_x^B] \rangle + \mathbb{V}[\langle m_{x|B} | \Sigma_x^B \rangle], \quad (124)$$

the first term being the mean noise variance and the second representing sample variance.

12. Where Should We Look for Extragalactic Broadcasts?

The idea that starfaring societies replicate until they pervade a galaxy suggests large galaxies as disproportionately favorable targets for SETI.

It is plausible that spreading metasocieties are very rare, with $\ll 1$ on average per Milky Way-sized galaxy. The lack of evidence for an omnipresent metasociety in our own Galaxy is suggestive, although perhaps that is the result of anthropic selection effects (Hanson et al. 2021) or active measures to remain hidden (Ball 1973). Rarity is also consistent with the

lack of evidence for Kardashev Type III metasocieties (Annis 1999; Garrett 2015; Griffith et al. 2015; Lacki 2016; Chen & Garrett 2021) and negative extragalactic SETI results (Horowitz & Sagan 1993; Shostak et al. 1996; Gray & Mooley 2017). With the framework developed, we now consider how many broadcasts to expect from galaxies of differing sizes, contrasting galactic clubs, and spreading metasocieties.

Suppose, under the interchangeability and independence assumptions, that societies in populated galaxies create ultra-narrowband beacons, all very bright and easily detected out to cosmological distances. In the line regime of the box model (or equivalently, the chord model with $\bar{\delta}_G = 0$), $\langle N_{s,C}^B \rangle = \check{Z}_C^B B_s$ for survey s .

What is the probability that s intercepts at least one such broadcast in a comoving volume? And do we expect them in the more numerous small galaxies, or in the rarer large galaxies? If these beacons are sufficiently rare, most observed societies have none active within the band observed by the survey ($\langle N_{s,C}^B \rangle \ll 1$), meaning Equation (102) is a good estimate of the probability that at least one broadcast passing the luminosity cut is intercepted. The stellar mass density distribution of galaxies with a detectable broadcast is given by

$$\frac{d^2 \langle N_G(N_{s,G}^B \geq 1) \rangle}{dV d \ln \langle M_{t,G}^* \rangle} = P(N_{s,G}^B \geq 1) \frac{d^2 \langle N_G \rangle}{dV d \ln \langle M_{t,G}^* \rangle}. \quad (125)$$

To find this mass function of galaxies with detected beacons, I adopt the $z \sim 0.1$ galaxy mass distributions from Moustakas et al. (2013), and assume $\langle m_* \rangle = 0.2 M_\odot$ motivated by Chabrier (2003).

I compare the expansive metasociety and galactic club scenario in Figure 9. Expansive metasocieties are assumed by default to be fairly rare ($\Xi_G^{M=0} = 10^{-12}$) but they plant one society around every star ($\Xi_M^C = 1$), so that $\langle N_{s,G}^C \rangle \approx 10^{10}$ for $\langle N_{t,G}^* \rangle = \langle N_{t,G}^C \rangle = 10^{11}$, a moderate sized galaxy. The comparable galactic club scenario has $\Xi_G^C = 0.1$, so that the mean number of societies in a galaxy of the same size is nearly equal, $\langle N_{s,G}^C \rangle = 10^{10}$ for $\langle N_{t,G}^* \rangle = 10^{11}$. This is likely extremely generous to the galactic club models because abundances this high require a mean societal lifespan of order 1 Gyr. Note that in the galactic club scenario, the societies are basically spread evenly between galaxies, while in the expansive metasociety scenario, only about 10% of $\langle N_{t,G}^* \rangle = 10^{11}$ galaxies are inhabited, but these have 10 times the mean number of societies.

As can be seen in Figure 9, there are two regimes of behavior: an asymptotic density distribution when $\langle N_{s,G}^B \rangle \gg \langle N_{t,G}^* \rangle^{-1}$, and a steeper form when $\langle N_{s,G}^B \rangle \ll \langle N_{t,G}^* \rangle^{-1}$. In the latter case, broadcasts are rare, and thus bigger galaxies are more likely to have one. Detections are biased to more massive galaxies in the expansive metasociety scenario, where $P(N_{s,G}^B \geq 1)$ has a quadratic dependence on the expected number of stars at the present, $\langle N_{t,G}^* \rangle$.

The reason for the quadratic dependence is simple: both $\langle N_{t,G}^M \rangle$ and $\langle N_{s,M}^C \rangle$ independently grow with stellar mass. When expansive metasocieties are rare, larger galaxies are proportionately more likely to host one simply because they have more stars for them to evolve around. Then, when one does appear and spread across the galaxy, there are more stars to populate, resulting in more societies, and thus more broadcasts. In contrast, while more stars mean more societies in the galactic

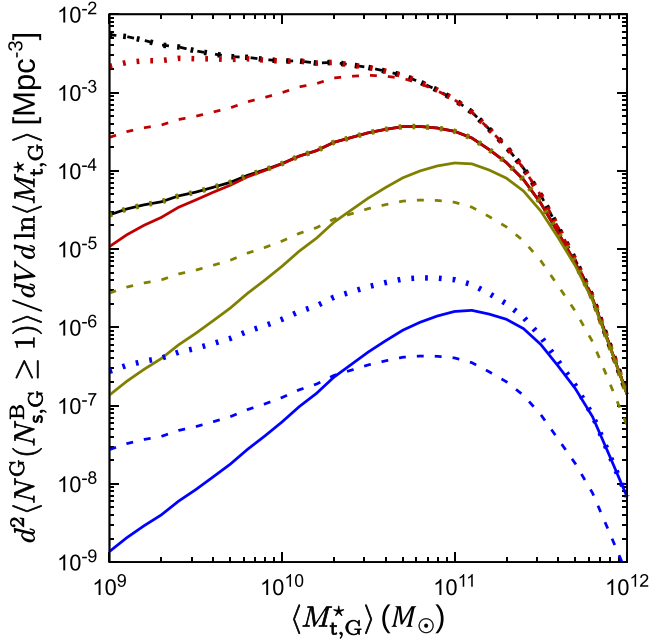


Figure 9. Number density of galaxies with at least one detected beacon in galactic clubs (dashed line, $\Xi_G^C = 0.1$) and expansive metasocieties (solid line for $\Xi_G^{M=0} = 10^{-12}$ and dotted line for $\Xi_G^{M=0} = 1$, both with $\Xi_M^C = 1$). Different colors are used for different $\langle N_{s,G}^B \rangle$: 100 (black), 10^{-10} (red), 10^{-12} (gold), and 10^{-14} (blue). The fraction of galaxies with detectable beacons can be near 1 (common broadcasts and societies), increase linearly with stellar mass (galactic clubs or common expansive metasocieties), or increase quadratically with stellar mass (rare expansive metasocieties).

club scenario, each society is confined to one star regardless of galaxy mass.⁴² This effect applies only when expansive metasocieties are rare; when they are extremely common (dotted lines in Figure 9), the linear dependence on mass is restored—though with a higher constant of proportionality because we expect many more societies in an expansive metasociety.

This demonstration suggests that if interstellar migration is a credible possibility, extragalactic SETI could do well to focus on high-mass galaxies. Breakthrough Listen’s nearby galaxy survey observes several large galaxies, including ellipticals in the Virgo cluster (Isaacson et al. 2017). It also has commensal access to MeerKAT, which will conduct deep observations of the Fornax cluster, which also is home to many large galaxies (Czech et al. 2021).

13. Conclusion

If interstellar travel and migration are indeed possible, then ETIs are unlike known astrophysical phenomena in that they can reproduce. Replication can amplify quirks of history onto galactic scales. Thus, supposing that starfaring ETIs are rare, one galaxy could have no ETIs while another, astrophysically indistinguishable, could have billions of inhabited worlds. This motivates the use of a probabilistic treatment of the observable technosignatures of galaxies, wherein different galaxies can have wildly different broadcast distributions.

⁴² There are additional reasons to favor larger galaxies: their stellar populations tend to be older (e.g., Gallazzi et al. 2005; Conroy et al. 2014), allowing for more time for ETIs to evolve, and they tend to have higher metallicities according to the mass–metallicity relationship, suggesting more planets may be around (Dayal et al. 2015).

This work introduces a framework that walks step by step from this population to measurements. Galaxies are treated as isolated systems within the universe. Within them, ETIs are congregated into localized societies. Entire lineages of societies are aggregated into metasocieties, which can cover the entire galaxy. Broadcasts are produced by societies. These objects are organized in a random tree structure, where objects can have ancestor hosts and in turn host descendants.

Each of these objects is described by a parameter tuple, points in their respective haystack (see Wright et al. 2018). Their populations are random point processes in the haystacks, one process defined for every possible host. Every point process has a distribution (intensity), describing the mean number of objects in different parts of the haystack. A selection is a thinning according to position in the haystack, within some window, for some host object. The result is a biased random set, which is realized as a sample. The Poisson point process is a simple description of populations in realized hosts when objects are independent and interchangeable. Observables can then be calculated using compound Poisson distributions. Given a distribution of random objects, the population is described by a mixture of Poisson point processes called a Cox point process. The box model and the chord models specifically describe the observables of broadcasts.

I present several examples and models to show the scope of the framework:

1. Lifespan bias, in which longer-lived objects have a larger temporal cross section, is viewed as the time-limited window reaching farther into the haystack as lifespan increases (Section 3.8; Figure 6). When sampling with a window x with duration T_x , the duration mean is biased as

$$\langle f(\tau) \rangle_{x,K} = \frac{\langle \tau f(\tau) \rangle_K + T_x \langle f(\tau) \rangle_K}{\langle \tau \rangle_K + T_x},$$

with each biased moment of object lifespan τ_j for short windows depending on the next unbiased moment. Selection bias also shows up in the chord model, where broadcasts with high drift rates are more likely to be sampled.

2. The formalism can be used to rederive the Drake equation for classical assumptions about the lack of significant interstellar travel (Section 6.4). It follows from integrating the societal and metasocietal distribution functions in the classical scenario. This work’s version says that the instantaneous mean number of communicative societies in a galaxy is

$$\langle N_{t,G}^C(t) \rangle \approx \tilde{f}_G^* \langle f_{\star}^M \rangle_G \langle f_M^C \rangle_G \langle \tau_M \rangle_G,$$

the product of a constant star formation rate, the mean fraction of stars hosting ETIs, the mean fraction of societies becoming communicative, and the mean lifespan of societies.

3. Expressions for the mean and variance of aggregate random variables are presented for *galactic club* and *expansive metasociety* scenarios are presented (Table 9), under the interchangeability assumption. They apply to the total intercepted emission from all broadcasts in a galaxy. The fluence of the population is proportional to the number of stars, rate/abundance of broadcasts, mean fluence of individual broadcasts, and additional factors

related to the bandwidth and duration of the broadcasts and windows (Section 8.4):

$$\langle \mathcal{U}_{x,j}^B \rangle \approx \Lambda_j^B \langle N_{x,j}^* \rangle (B_x + \langle \beta_{x;l(B)} \rangle) (\tau_x + \langle \tau_B \rangle) \cdot \langle u_{x|B} \rangle_j.$$

Variance in the aggregate variable increases as the interchangeability assumption is relaxed (Section 8.6).

4. The presence of expansive metasocieties can actually result in the number of detected broadcasts depending quadratically on stellar mass (Section 12). This is because the number of opportunities for an interstellar metasociety to arise in the first place is proportional to the number of stars, and then independently, the number of sites for independent broadcasting societies is again proportional to the number of stars. Larger galaxies may then be disproportionately likely to host detectable technosignatures (Figure 9).

Future papers of the series will consider limits on broadcast populations from total radio emission and individual searches in the face of confusion (Paper II), as well as constraints on populations of broadcasting galaxies from source counts and commensal searches of background galaxies (Paper III).

The ideas behind the framework can be applied to other phenomena, though with a different hierarchy of objects. Transients like fast radio bursts are generated by discrete objects like neutron stars, each of which could yield a whole population of events. In turn, some kinds of sites may be found only in globular clusters or other subgalactic environments (e.g., Kirsten et al. 2022). Thus, like broadcasts, natural transients can be clustered into multiple levels of hosts, which may be described by point processes.

Acknowledgments

I thank the Breakthrough Listen program for their support. Funding for Breakthrough Listen research is sponsored by the Breakthrough Prize Foundation (<https://breakthroughprize.org/>), which manages Breakthrough Initiatives. In addition, I acknowledge the use of NASA's Astrophysics Data System and arXiv for this research. I am grateful to the referee for the dedicated work in reviewing Papers I and II, including thorough readings; and I thank the statistics editor for guiding me to useful references.

Appendix A

Extreme Value Theory and Regularization

A.1. Review of Basic Extreme Value Theory

Suppose we have a collection of n random variables X_i (with i ranging from 1 to n), all i.i.d. We can sort the values assigned to the X_i and define the order statistic $X_{(i)}$ as the i th smallest value. The order statistics themselves are random variables, each with their own well-defined probability distribution (e.g., Gumbel 1958; Coles et al. 2001; Castillo et al. 2005; Embrechts et al. 2013). Most noteworthy are the minimum $X_{(1)} = \min X_i$ and the maximum $X_{(n)} = \max X_i$. Their probability distributions are easy to calculate since it is just the probability that every X_i independently is greater or smaller than x . Given the CDF $F[X]$ and the CCDF $\bar{F}[X] \equiv 1 - F[X]$, the minimum's CDF is

$$F[X_{(n)}|n](x) = \prod_{i=1}^n F[X_i](x) = [F[X_i](x)]^n \quad (A1)$$

and the minimum's CCDF is

$$\bar{F}[X_{(1)}](x) = \prod_{i=1}^n \bar{F}[X_i](x) = [\bar{F}[X_i](x)]^n. \quad (A2)$$

These random variables generally converge to one of the three types of extreme value distributions—the Gumbel distribution if the X_i has an exponential-like PDF tail (including the normal distribution), the Fréchet distribution when X_i has a power-law-like PDF tail, and the Weibull distribution when they have finite support (e.g., Gumbel 1958; Castillo et al. 2005).

However, when observing a population of astronomical objects in a field, the actual number sampled is a random variable N . For Poissonian N , $X_{(1)} = \min_{1 \leq i \leq N} X_i$ and $X_{(N)} = \max_{1 \leq i \leq N} X_i$ each have a mixture probability distribution, a weighted sum of the probability distributions conditionalized on each value of n :

$$F[X_{(N)}](x) = \frac{\sum_{n=1}^{\infty} P(N=n) F[X_{(N)}|N](x|n)}{1 - P(N=0)} = \frac{e^{\langle N \rangle F[X](x)} - 1}{e^{\langle N \rangle} - 1}$$

$$\bar{F}[X_{(1)}](x) = \frac{\sum_{n=1}^{\infty} P(N=n) \bar{F}[X_{(1)}|N](x|n)}{1 - P(N=0)} = \frac{e^{\langle N \rangle \bar{F}[X](x)} - 1}{e^{\langle N \rangle} - 1} \quad (A3)$$

(Castillo et al. 2005). Note that the case where $N=0$ is specifically excluded because the minimum and maximum X_i in a sample with zero members is not defined. This exclusion is necessary to ensure the minimum and maximum exist.

A.2. Regularization Using Median Values for Minimum and Maximum Values in a Poissonian Sample

In this series, regularization is accomplished by introducing cutoffs in the values of X . I use quartiles for the minimum and maximum X in a sample, with a lower bound given by $\bar{F}[X_{(1)}](X^L) = 3/4$ and an upper bound of $F[X_{(N)}](X^H) = 3/4$. Equation (A3) gives us

$$F[X](X^H) = \frac{1}{\langle N \rangle} \ln \left[\frac{3}{4} e^{\langle N \rangle} + \frac{1}{4} \right]$$

$$F[X](X^L) = 1 - \frac{1}{\langle N \rangle} \ln \left[\frac{3}{4} e^{\langle N \rangle} + \frac{1}{4} \right]. \quad (A4)$$

When $\langle N \rangle \gg 1$, the great majority of the probability mass is sampled. As long as we have the inverse CDF of X , the bounds are approximated as $X^H \approx F[X]^{-1}(1 + \ln(3/4)/\langle N \rangle)$ and $X^L \approx F[X]^{-1}(-\ln(3/4)/\langle N \rangle)$.

When $0 < \langle N \rangle \ll 1$, $F[X](X^L)$ and $F[X](X^H)$ converge to one-quarter and three-quarters, respectively. Roughly, the spread among all X_i should be the spread in a single X_i because virtually all observations have no events and almost all of the rest have $N=1$. We could also use other probability thresholds for X^L and X^H . Using the median $X_{(1)}$ and $X_{(N)}$ is an intuitive choice when $\langle N \rangle \gtrsim 1$, but for $\langle N \rangle \ll 1$, it would imply that the regularized variable has a variance approaching zero, and is thus unsatisfactory.

Appendix B

Further Details of the Box Model

The distributions and averages I present here are calculated assuming a uniform spread in time and frequency (uniform Λ_j^B) and that $(\bar{\beta} + B_x)/2 \leq \gamma_x$.

B.1. The Overlap Fraction

If we regard time and frequency as a plane, the overlap fraction $o_{x|B}$ is the area of intersection between two contiguous boxes, scaled to the area of the broadcast's box:

$$o_{x|B} \equiv \frac{\tau_{x|B}\beta_{x|B}}{\bar{\tau}\bar{\beta}} = \frac{R[\min(\vartheta_B + \bar{\tau}, \Theta_x + \bar{\tau}_x) - \max(\vartheta_B, \Theta_x)]}{\bar{\tau}} \cdot \frac{R\left[\min\left(v_B + \frac{\bar{\beta}}{2}, \Upsilon_x + \frac{B_x}{2}\right) - \max\left(v_B - \frac{\bar{\beta}}{2}, \Upsilon_x - \frac{B_x}{2}\right)\right]}{\bar{\beta}}, \quad (B1)$$

using the ramp function $R(x) = \max(0, x)$. The overlap can be decomposed into the intersection between the time ranges spanned by the sample window and the broadcast and the intersection of the frequency ranges spanned by the sample and broadcast. It reaches a maximum value when one of the time ranges is fully inside the other and likewise, one frequency range contains the other, although it is possible for the sample window to contain the broadcast for the one and the broadcast to contain the window for the other. This maximal value is

$$o_{x|B}^{\max} = \min\left(1, \frac{\bar{\tau}_x}{\bar{\tau}}\right) \min\left(1, \frac{B_x}{\bar{\beta}}\right). \quad (B2)$$

Among broadcasts within the sample (with a nonzero $o_{x|B}$), the probability this value is attained is

$$P(o_{x|B} = o_{x|B}^{\max}) = \frac{|\bar{\tau} - \bar{\tau}_x||\bar{\beta} - B_x|}{(\bar{\tau} + \bar{\tau}_x)(\bar{\beta} + B_x)}. \quad (B3)$$

This probability only diverges significantly from 1 if $\bar{\tau} \approx \bar{\tau}_x$ or $\bar{\beta} \approx B_x$. A priori we may expect that ETIs and our own efforts are uncoordinated and thus this is unlikely to happen, but there are natural lower limits to bandwidth set by interstellar scattering (Cordes et al. 1997), and thus our observations may try to match the broadened bandwidth of a line. In any case, only in rare cases will the partial overlap result in $0 < o_{x|B} \ll o_{x|B}^{\max}$. The probability density for this regime is

$$\psi[o_{x|B}]_x(o|0 < o < o_{x|B}^{\max}) = \frac{2}{(\bar{\tau} + \bar{\tau}_x)(\bar{\beta} + B_x)} \left[\max\left(\frac{\bar{\tau}_x}{\bar{\tau}}, \frac{\bar{\tau}}{\bar{\tau}_x}\right) + \max\left(\frac{B_x}{\bar{\beta}}, \frac{\bar{\beta}}{B_x}\right) - 2 - 2 \ln\left(\frac{o}{o_{x|B}^{\max}}\right) \right]. \quad (B4)$$

The means used in the next section can be calculated from Equations (B2)–(B4).

B.2. Emission and Selection Windows

The emission that falls into the selection box is equal to the total emission ever released in the broadcast multiplied by the overlap fraction and the fraction in observed polarizations: $\hat{\epsilon}_{x|B} = \hat{\epsilon}_B o_{x|B} \zeta_{x|B}$. Thus, the mean restricted energy is

$$\langle \hat{\epsilon}_{x|B} \rangle_J = \frac{|\Pi_x|}{2} \langle \hat{\epsilon}_B \rangle_{x,J} \frac{\bar{\tau}_x B_x}{(\bar{\tau}_x + \bar{\tau})(B_x + \bar{\beta})}. \quad (B5)$$

The variance is found using

$$\langle \hat{\epsilon}_{x|B}^2 \rangle_J = \langle \zeta_{x|B}^2 \rangle_J \langle \hat{\epsilon}_B^2 \rangle_{x,J} \frac{\bar{\tau}_x B_x}{(\bar{\tau}_x + \bar{\tau})(B_x + \bar{\beta})} \max\left[1 - \frac{\bar{\tau}}{3\bar{\tau}_x}, \frac{\bar{\tau}_x}{\bar{\tau}} - \frac{\bar{\tau}_x^2}{3\bar{\tau}^2}\right] \max\left[1 - \frac{\bar{\beta}}{3B_x}, \frac{B_x}{\bar{\beta}} - \frac{B_x^2}{3\bar{\beta}^2}\right]. \quad (B6)$$

Equation (113) implies the number of photons per unit frequency falls as ν^{-1} . I find

$$\langle \hat{q}_{x|B} \rangle_J = \frac{|\Pi_x|}{2} \frac{\langle \hat{\epsilon}_B \rangle_{x,J}}{h} \frac{\bar{\tau}_x}{(\bar{\tau}_x + \bar{\tau})(B_x + \bar{\beta})} \ln\left(\frac{1 + B_x/(2\bar{\tau}_x)}{1 - B_x/(2\bar{\tau}_x)}\right), \quad (B7)$$

approaching $\langle \hat{\epsilon}_{x|B} \rangle / (h\bar{\tau}_x)$ when $B_x \ll \bar{\tau}_x$. Furthermore,

$$\langle \dot{q}_{x|B}^2 \rangle_J = \langle \zeta_{x|B}^2 \rangle_J \frac{\langle \dot{\epsilon}_B^2 \rangle_{x,J}}{h} \frac{T_x}{(T_x + \bar{\tau})(B_x + \bar{\beta})} \max \left[1 - \frac{\bar{\tau}}{3T_x}, \frac{T_x}{\bar{\tau}} - \frac{T_x^2}{3\bar{\tau}^2} \right] \frac{1}{\bar{\beta}^2} \cdot \begin{cases} \left[4B_x + 4T_x \ln \frac{1 - B_x/(2T_x)}{1 + B_x/(2T_x)} + \bar{\beta} \left[\ln \frac{1 - B_x/(2T_x)}{1 + B_x/(2T_x)} \right]^2 \right] & \text{if } \bar{\beta} \geq B_x \\ 2 \left[\bar{\beta} \left(2 + \text{Li}_2 \frac{\bar{\beta}}{T_x - B_x/2 + \bar{\beta}} - \text{Li}_2 \frac{\bar{\beta}}{T_x + B_x/2} - \frac{1}{2} \left[\ln \left(1 + \frac{\bar{\beta}}{T_x - \bar{\beta}/2} \right) \right]^2 \right) \right. \\ \quad \left. + (T_x - B_x/2 + \bar{\beta}) \ln \left(1 + \frac{\bar{\beta}}{T_x - B_x/2} \right) \left[\ln \left(1 + \frac{\bar{\beta}}{T_x - B_x/2} \right) - 1 \right] \right. \\ \quad \left. - (T_x + B_x/2 - \bar{\beta}) \ln \left(1 - \frac{\bar{\beta}}{T_x + B_x/2} \right) \left[\ln \left(1 - \frac{\bar{\beta}}{T_x + B_x/2} \right) - 1 \right] \right] & \text{if } \bar{\beta} \leq B_x \end{cases}, \quad (\text{B8})$$

with $\text{Li}_2(x)$ being the dilogarithm (Spence's function). If $B_x \ll T_x$, $\langle \dot{q}_{x|B}^2 \rangle_J \approx \langle \zeta_{x|B}^2 \rangle_J \langle \dot{\epsilon}_{x|B}^2 \rangle_J / (hT_x)^2$.

Mean fluences may be found by multiplying by the mean of the appropriate transmittance/dilution factor, $\langle \chi_{\rho;B}^n y_{\rho;B}^n \rangle_{x,J}$ for $\langle \dot{\epsilon}_{x|B}^n \rangle_{x,J}$ (Equation (91)).

Approximations when $\bar{\tau} \approx T_x$ and $\bar{\beta} \approx B_x$ are given in Table 10.

Appendix C

Further Details of the Chord Model

The results presented here assume a constant Z_J^B .

C.1. Selection of Chords in a Contiguous Selection Window

In the chord model, the window on selected times and frequency is a contiguous box, as in the box model. Because we only constrain the behavior of the broadcast within the sample window, calculations are simplified when we cut the chord, only considering its behavior at times between Θ_x and $\Theta_x + T_x$. This imposes $\vartheta_{x:t|B} = \Theta_x$, $\tau_{x:t|B} = T_x$, and $\beta_{x:t|B} = |\delta_B|T_x$. I also use $v_{x:t|B}$, the frequency during the mid-time of the window. At other times, the band is centered on a frequency $v_{t|B}(t) = v_{x:t|B} + (t - \Theta_x - T_x/2)\delta_B$. At any given frequency, the center of the band crosses at time $\vartheta_{\nu|B}(\nu) = (\Theta_x + T_x/2) + (\nu - v_{x:t|B})/\delta_B$. Note that $T_x = T_x^\oplus/(1+z)$, $B_x = B_x^\oplus/(1+z)$, and $\delta_B = \delta_B^\oplus/(1+z)^2$ are in the source frame.

A broadcast crosses the selection window if $|\vartheta_{x:t|B} - v_{x:t|B}| \leq (B_x + T_x|\delta_B|)/2$. The mean number of broadcasts intercepted is thus given by

$$\frac{d\langle N_{x,J}^B \rangle}{d\delta_B} = \langle N_{x,J}^* \rangle Z_J^B (B_x + T_x|\delta_B|) \psi[\delta_B]_J, \quad (\text{C1})$$

noting that $\psi[\delta_B]_J$ is the unbiased drift rate distribution. It immediately follows that $\langle N_{x,J}^B \rangle = \langle N_{x,J}^* \rangle Z_J^B (B_x + T_x\langle |\delta_B| \rangle_J)$, regardless of distribution, where $\langle |\delta_B| \rangle_J$ is the unbiased mean drift rate magnitude. The biased drift rate distribution, that is, the distribution sampled by the window, favors high drift rate broadcasts:

$$\psi[\delta_B]_{x,J}(\delta) = \frac{1}{\langle N_{x,J}^B \rangle} \frac{d\langle N_{x,J}^B \rangle}{d\delta_B}(\delta) = \frac{B_x + T_x|\delta|}{B_x + T_x\langle |\delta_B| \rangle_J} \psi[\delta_B]_J(\delta), \quad (\text{C2})$$

(see Equation (36)).

When expressing calculations, it is convenient to define dimensionless variables, placing the drift rates in *natural* units defined by the window: $u_{x;B} \equiv \delta_B T_x / B_x$ and $U_{x;J} \equiv \Delta_J^B T_x / B_x$. Chords with $|u_{x;B}| \ll 1$ behave like lines in the box model.

According to the uniform drift rate distribution model, the drift rates have a uniform unbiased distribution (Equation (120)) spanning the range given by $|\Delta_J^B - \delta_B| \leq \bar{\delta}_J$. We have

$$\langle |\delta_B| \rangle_J = \begin{cases} \bar{\delta}_J/2 \cdot (1 + (\Delta_J^B/\bar{\delta}_J)^2) & \text{if } |\Delta_J^B| \leq \bar{\delta}_J \\ |\Delta_J^B| & \text{if } |\Delta_J^B| \geq \bar{\delta}_J \end{cases} \quad (\text{C3})$$

and $\langle \delta_B^2 \rangle_J = \bar{\delta}_J^2/3 + \Delta_J^B$. The mean drift rate of a sampled broadcast comes from Equation (C2):

$$\langle |\delta_B| \rangle_{x,J} = \begin{cases} \frac{\bar{\delta}_J}{2} \frac{1 + (2/3)\bar{u}_{x,J} + 2(U_{x,J})^2 / \bar{u}_{x,J} + (U_{x,J} / \bar{u}_{x,J})^2}{1 + (1/2)(\bar{u}_{x,J} + (U_{x,J})^2 / \bar{u}_{x,J})} & \text{if } |\Delta_J^B| \leq \bar{\delta}_J \\ \frac{B_x}{T_x} \frac{(1/3)\bar{u}_{x,J}^2 + |U_{x,J}| + (U_{x,J})^2}{1 + |U_{x,J}|} & \text{if } |\Delta_J^B| \geq \bar{\delta}_J, \end{cases} \quad (C4)$$

with $\bar{u}_{x,J} \equiv \bar{\delta}_J T_x / B_x$.

Furthermore, although the unbiased drift rate magnitude CDF is $F[|\delta_B|_J](\delta) = F[\delta_B]_J(|\delta|) - F[\delta_B]_J(-|\delta|)$, the biased CDF is

$$F[|\delta_B|]_{x,J}(u_{x,B}) = \begin{cases} \frac{u_{x,B}}{\bar{u}_{x,J}} \frac{1 + (1/2)u_{x,B}}{1 + (1/2)(\bar{u}_{x,J} + (U_{x,J})^2 / \bar{u}_{x,J})} & \text{if } |\Delta_J^B| \leq \bar{\delta}_J \text{ and } \bar{\delta}_J - |\Delta_J^B| \geq \delta \\ \frac{u_{x,B}^2 + 2u_{x,B} + 2(\bar{u}_{x,J} - |U_{x,J}|) + (\bar{u}_{x,J} - |U_{x,J}|)^2}{4u_{x,B}[1 + (1/2)(\bar{u}_{x,J} + (U_{x,J})^2 / \bar{u}_{x,J})]} & \text{if } |\Delta_J^B| \leq \bar{\delta}_J \text{ and } \bar{\delta}_J - |\Delta_J^B| \leq \delta \leq \bar{\delta}_J + |\Delta_J^B| \\ 0 & \text{if } |\Delta_J^B| \geq \bar{\delta}_J \text{ and } |\Delta_J^B| - \bar{\delta}_J \geq \delta \\ \frac{u_{x,B}^2 + 2u_{x,B} - 2(|U_{x,J}| - \bar{u}_{x,J}) - (|U_{x,J}| - \bar{u}_{x,J})^2}{4u_{x,B}[1 + |U_{x,J}|]} & \text{if } |\Delta_J^B| \geq \bar{\delta}_J \text{ and } |\Delta_J^B| - \bar{\delta}_J \leq \delta \leq |\Delta_J^B| + \bar{\delta}_J \\ 1 & \text{if } \bar{\delta}_J + |\Delta_J^B| \leq \delta, \end{cases} \quad (C5)$$

with $u_{x,B} \equiv \delta_B T_x / B_x$. When the drift rates span a large range ($\bar{\delta}_J \gg B_x / T_x$ and $\geq |\Delta_J^B|$), covering the entire range of integration, the biased probability has a quadratic dependence on $\delta / \bar{\delta}_J$.

C.2. Statistics of Chord Duration in Sample

The central quantity in the chord model is $\tau_{x|B}$, the time it takes the chord to cross the window:

$$\tau_{x|B} = \min[\vartheta_{x|B}(\Upsilon_x + \text{sgn}(\delta_B)B_x/2), \Theta_x + T_x] - \max[\vartheta_{x|B}(\Upsilon_x - \text{sgn}(\delta_B)B_x/2), \Theta_x]. \quad (C6)$$

It can be shown that

$$\tau_{x|B} = \begin{cases} \min\left(\frac{B_x}{|\delta_B|}, T_x\right) & \text{if } |\Upsilon_x - v_{x:tlB}| \leq \frac{|\Upsilon_x| \delta_B - B_x}{2} \\ \frac{T_x + B_x / |\delta_B|}{2} - \frac{|\Upsilon_x - v_{x:tlB}|}{|\delta_B|} & \text{if } \frac{|\Upsilon_x| \delta_B - B_x}{2} \leq |\Upsilon_x - v_{x:tlB}| \leq \frac{|\Upsilon_x| \delta_B + B_x}{2} \\ 0 & \text{if } |\Upsilon_x - v_{x:tlB}| \geq \frac{|\Upsilon_x| \delta_B + B_x}{2}. \end{cases} \quad (C7)$$

For a population in a host J with a uniform v_B distribution, sampled broadcasts with a fixed drift rate (necessarily with nonzero $\tau_{x|B}$) have

$$\langle \tau_{x|B}^n |\delta_B \rangle_{x,J} = \frac{T_x^n}{1 + |u_{x,B}|} \begin{cases} 1 - \frac{n-1}{n+1} |u_{x,B}| & \text{if } |u_{x,B}| \leq 1 \text{ and } n \neq -1 \\ \frac{1}{|u_{x,B}|^n} \left(|u_{x,B}| - \frac{n-1}{n+1} \right) & \text{if } |u_{x,B}| \geq 1 \text{ and } n \neq -1. \end{cases} \quad (C8)$$

The mean of a random variable for a broadcast sample is weighted by the number of broadcasts that are in the sample. This leads to the biasing of the mean toward high drift rate broadcasts:

$$\langle f(\tau_{x|B}) \rangle_{x,J} = \int_{-\infty}^{\infty} \langle f(\tau_{x|B}) |\delta_B \rangle_{x,J} \psi[\delta_B]_{x,J} d\delta_B = \frac{1}{B_x + T_x \langle |\delta_B| \rangle_J} \int_{-\infty}^{\infty} (B_x + T_x |\delta_B|) \langle f(\tau_{x|B}) |\delta_B \rangle \psi[\delta_B]_J d\delta_B \quad (C9)$$

for the random variable defined by applying the function f to $\tau_{x|B}$. This immediately gives us the x-relative mean

$$\langle \tau_{x|B} \rangle_J = \frac{T_x}{1 + T_x \langle |\delta_B| \rangle_J / B_x}. \quad (C10)$$

For the uniform drift rate distribution, I find that

$$\langle \tau_{x|B}^2 \rangle_J = T_x \langle \tau_{x|B} \rangle_J \begin{cases} 1 - (1/6)(\bar{u}_{x,J} + (U_{x,J})^2/\bar{u}_{x,J}) & \text{if } |U_{x,J}| \leq \bar{u}_{x,J} \text{ and } \bar{u}_{x,J} + |U_{x,J}| \leq 1 \\ \frac{1}{4\bar{u}_{x,J}} \left[1 + 2(\bar{u}_{x,J} - |U_{x,J}|) - (1/3)(\bar{u}_{x,J} - |U_{x,J}|)^2 + 2\ln(\bar{u}_{x,J} + |U_{x,J}|) + \frac{(2/3)}{\bar{u}_{x,J} + |U_{x,J}|} \right] & \text{if } |U_{x,J}| \leq \bar{u}_{x,J} \text{ and } |1 - \bar{u}_{x,J}| \leq |U_{x,J}| \\ \frac{1}{2\bar{u}_{x,J}} \left[1 + \ln(\bar{u}_{x,J}^2 - |U_{x,J}|^2) + \frac{(1/3)}{\bar{u}_{x,J} + |U_{x,J}|} + \frac{(1/3)}{\bar{u}_{x,J} - |U_{x,J}|} \right] & \text{if } |U_{x,J}| \leq \bar{u}_{x,J} \text{ and } 1 \leq \bar{u}_{x,J} - |U_{x,J}| \\ 1 - (1/3)|U_{x,J}| & \text{if } |U_{x,J}| \geq \bar{u}_{x,J} \text{ and } \bar{u}_{x,J} + |U_{x,J}| \leq 1 \\ \frac{1}{4\bar{u}_{x,J}} \left[1 - 2(|U_{x,J}| - \bar{u}_{x,J}) + (1/3)(|U_{x,J}| - \bar{u}_{x,J})^2 + 2\ln(\bar{u}_{x,J} + |U_{x,J}|) + \frac{(2/3)}{\bar{u}_{x,J} + |U_{x,J}|} \right] & \text{if } |U_{x,J}| \geq \bar{u}_{x,J} \text{ and } |1 - |U_{x,J}|| \leq |\bar{u}_{x,J}| \\ \frac{1}{2\bar{u}_{x,J}} \left[\ln \frac{|U_{x,J}| + \bar{u}_{x,J}}{|U_{x,J}| - \bar{u}_{x,J}} + \frac{(1/3)}{|U_{x,J}| + \bar{u}_{x,J}} - \frac{(1/3)}{|U_{x,J}| - \bar{u}_{x,J}} \right] & \text{if } |U_{x,J}| \geq \bar{u}_{x,J} \text{ and } 1 \leq |U_{x,J}| - \bar{u}_{x,J}. \end{cases} \quad (C11)$$

C.3. Emission and Fluence

The effective isotropic energy is found using the (source-frame) effective luminosity and the (source-frame) chord crossing time:

$$\dot{\epsilon}_{x|B} = \zeta_{x|B} \dot{\ell}_B \tau_{x|B}. \quad (C12)$$

The polarization, luminosity, and chord duration are all presumed independent, so the mean energy fluence is

$$\langle \dot{\epsilon}_{x|B} \rangle_J = \frac{|\Pi_x|}{2} \langle \dot{\ell}_B \rangle_{x,J} \frac{T_x}{1 + T_x \langle |\delta_B| \rangle_J / B_x}. \quad (C13)$$

If the uniform drift rate distribution is adopted and the mean drift rate is near zero ($|\Delta_B^J| \ll \bar{\delta}_J$),

$$\langle \dot{\epsilon}_{x|B}^2 \rangle_J = \langle \zeta_{x|B}^2 \rangle_J \langle \dot{\ell}_B^2 \rangle_{x,J} \langle \tau_{x|B}^2 \rangle_J \approx \langle \zeta_{x|B}^2 \rangle_J \langle \dot{\ell}_B^2 \rangle_{x,J} T_x^2 \cdot \begin{cases} \frac{1 - (1/6)\bar{u}_{x,J}}{1 + (1/2)\bar{u}_{x,J}} & \text{if } \bar{u}_{x,J} \leq 1 \\ \frac{1}{2\bar{u}_{x,J}} \frac{1 + \ln \bar{u}_{x,J} + (2/3)\bar{u}_{x,J}^{-1}}{1 + (1/2)\bar{u}_{x,J}} & \text{if } \bar{u}_{x,J} \geq 1. \end{cases} \quad (C14)$$

From this windowed energy, it is easy to find $\langle h_{x|B} \rangle_J$ (Equation (91)):

$$\langle h_{x|B} \rangle_J = \frac{|\Pi_x|}{2} \langle \dot{\ell}_B \rangle_{x,J} \langle \chi_{\epsilon;x|B} \mathcal{Y}_{\epsilon;B} \rangle_{x,J} \frac{T_x}{1 + T_x \langle |\delta_B| \rangle_J / B_x}. \quad (C15)$$

In the chord model, if $\bar{\delta}_J \gg B_x/T_x$, the typical broadcast in a sample is a high drift rate signal that slashes through the sample box from low to high frequency or vice versa. Fixing broadcast effective isotropic luminosity, these typical broadcasts are much fainter than the much slower drifting broadcasts with $\delta_B \approx 0$. Now, the mean total fluence is found by integrating the sampled drift rate distribution,

$$\frac{d\langle \mathfrak{H}_{x,J}^B \rangle}{d\delta_B} = \frac{d\langle N_{x,J}^B \rangle}{d\delta_B} \langle h_{x|B} | \delta_B \rangle_{x,J} = \frac{|\Pi_x|}{2} \langle \dot{\ell}_B \rangle_J \langle \chi_{\epsilon;x|B} \mathcal{Y}_{\epsilon;B} \rangle_{x,J} Z_J^B \langle N_{x,J}^* \rangle T_x B_x \psi[\delta_B]_J. \quad (C16)$$

We find $\langle \mathfrak{H}_{x,J}^B (|\delta_B| \leq \delta) \rangle = \langle \mathfrak{H}_{x,J}^B \rangle F[|\delta_B|]_J(\delta)$ (note the use of the unbiased CDF). The fastest drifting half of the *population* contributes half the expected fluence. Yet this half is overrepresented in a sample. In the uniform drift rate model, if $|\delta_B| \gg B_x/T_x$, the bottom quartile of *sampled* broadcasts in $|\delta_B|$ contribute half the fluence.

As a result, the observed $\mathfrak{H}_{x,J}^B$ typically underestimates $\langle \mathfrak{H}_{x,J}^B \rangle$ because small samples tend to miss the slowest drifting broadcasts. For a uniform drift rate distribution, a few broadcasts suffice to counteract this effect. Biasing becomes more of an issue if the distribution has a long tail. For an exponential distribution $\psi[\delta_B]_J = (1/\bar{\delta}_J) \exp(-\delta_B/\bar{\delta}_J)$, the fraction of sampled broadcasts that contribute half the fluence falls to $(1 - \ln 2)/2 \approx 15\%$ as $\bar{\delta}_J \rightarrow \infty$. Thus, we might need to sample ~ 5 – 10 broadcasts until $\mathfrak{H}_{x,J}^B$ starts converging to $\langle \mathfrak{H}_{x,J}^B \rangle$. Heavy-tailed distributions like power laws would require still more broadcasts to be sampled before the aggregate emission converges to expectations from $\langle \mathfrak{H}_{x,J}^B \rangle$.

ORCID iDs

Brian C. Lacki  <https://orcid.org/0000-0003-1515-4857>

References

- Abbot, D. S., & Switzer, E. R. 2011, *ApJL*, **735**, L27
Adams, F. C., & Laughlin, G. 1997, *RvMP*, **69**, 337

- Adelson, R. M. 1966, *J. Oper. Res. Soc.*, **17**, 73
Annis, J. 1999, *JBIS*, **52**, 19
Armstrong, S., & Sandberg, A. 2013, *AcAau*, **89**, 1
Arnold, L. F. A. 2005, *ApJ*, **627**, 534
Baddeley, A. 2007, in *Stochastic Geometry*, ed. W. Weil (Berlin: Springer), 1
Badescu, V. 2011, *Icar*, **216**, 485
Badescu, V., & Cathcart, R. B. 2006, *AcAau*, **58**, 119
Bailyn, C. D. 1995, *ARA&A*, **33**, 133
Balbi, A., & Ćirković, M. M. 2021, *AJ*, **161**, 222

- Balbi, A., & Tombesi, F. 2017, *NatSR*, **7**, 16626
- Ball, J. A. 1973, *Icar*, **19**, 347
- Barbour, A. D., & Chryssaphinou, O. 2001, *Ann. Appl. Probab.*, **11**, 964
- Bas, E. 2019, *Basics of Probability and Stochastic Processes* (Berlin: Springer)
- Benford, J. N., & Benford, D. J. 2016, *ApJ*, **825**, 101
- Blair, D. G., & Zadnik, M. G. 1993, *A&A*, **278**, 669
- Borra, E. F. 2012, *AJ*, **144**, 181
- Bracewell, R. N. 1975, *The Galactic Club: Intelligent Life in Outer Space* (San Francisco, CA: Freeman)
- Brillinger, D. R. 1969, *Ann. Inst. Stat. Math.*, **21**, 215
- Brin, G. D. 1983, *QJRAS*, **24**, 283
- Carrigan, R. A. 2012, *AcAau*, **78**, 121
- Carroll-Nellenback, J., Frank, A., Wright, J., & Scharf, C. 2019, *AJ*, **158**, 117
- Carter, B. 1983, *RSPTA*, **310**, 347
- Castillo, E., Hadi, A. S., Balakrishnan, N., & Sarabia, J. M. 2005, *Extreme Value and Related Models with Applications in Engineering and Science* (New York: Interscience (Wiley-Interscience))
- Chabrier, G. 2003, *ApJL*, **586**, L133
- Chen, H., & Garrett, M. A. 2021, *MNRAS*, **507**, 3761
- Chennamangalam, J., Siemion, A. P. V., Lorimer, D. R., & Werthimer, D. 2015, *NewA*, **34**, 245
- Chiu, S. N., Stoyan, D., Kendall, W. S., & Mecke, J. 2013, *Stochastic Geometry and its Applications* (3rd ed.; New York: Wiley)
- Ćirković, M. M. 2018a, *AcAau*, **152**, 289
- Ćirković, M. M. 2018b, *The Great Silence: Science and Philosophy of Fermi's Paradox* (New York: Oxford Univ. Press)
- Ćirković, M. M., & Bradbury, R. J. 2006, *NewA*, **11**, 628
- Ćirković, M. M., & Vukotić, B. 2008, *OLEB*, **38**, 535
- Ćirković, M. M., & Vukotić, B. 2020, arXiv:2007.12645
- Coles, S., Bawa, J., Trenner, L., & Dorazio, P. 2001, *An Introduction to Statistical Modeling of Extreme Values*, Vol. 208 (Berlin: Springer)
- Conroy, C., Graves, G. J., & van Dokkum, P. G. 2014, *ApJ*, **780**, 33
- Corbet, R. H. D. 1997, *JBIS*, **50**, 253
- Corbet, R. H. D. 2003, *AsBio*, **3**, 305
- Cordes, J. M., Lazio, J. W., & Sagan, C. 1997, *ApJ*, **487**, 782
- Crick, F. H. C., & Orgel, L. E. 1973, *Icar*, **19**, 341
- Czech, D., Isaacson, H., Pearce, L., et al. 2021, *PASP*, **133**, 064502
- Daley, D. J., & Vere-Jones, D. 2003, *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods* (New York: Springer)
- Daley, D. J., & Vere-Jones, D. 2008, *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure* (New York: Springer)
- Davies, P. C. W., & Wagner, R. V. 2013, *AcAau*, **89**, 261
- Dayal, P., Cockell, C., Rice, K., & Mazumdar, A. 2015, *ApJL*, **810**, L2
- Di Stefano, R., & Ray, A. 2016, *ApJ*, **827**, 54
- Djorgovski, S. G., Mahabal, A., Drake, A., Graham, M., & Donalek, C. 2013, in *Planets, Stars and Stellar Systems. Volume 2: Astronomical Techniques, Software and Data*, ed. T. D. Oswalt & H. E. Bond (Berlin: Springer), 223
- Drake, F. D. 1961, *PtT*, **14**, 40
- Drake, F. D. 1965, in *Current Aspects of Exobiology*, ed. G. Mamikunian & M. H. Briggs (Oxford: Pergamon), 323
- Drake, F. D., Kardashev, N. S., Troitsky, V. S., et al. 1973, in *Communication With Extraterrestrial Intelligence*, ed. C. Sagan (Cambridge, MA: MIT Press), 230
- Drake, F. D., & Sagan, C. 1973, *Natur*, **245**, 257
- Dreher, J. W. 2004, in *IAU Symp. 213, Bioastronomy 2002: Life among the Stars*, ed. R. Norris & F. Stootman (San Francisco, CA: ASP), 467, IAU Symp. 213, Bioastronomy 2002: Life among the Stars
- Dyson, F. J. 1960, *Sci*, **131**, 1667
- Dyson, F. J. 1963, in *Interstellar Communication*, ed. A. G. W. Cameron (New York: Benjamin), 115
- Embrechts, P., Klüppelberg, C., & Mikosch, T. 2013, *Modeling Extremal Events: For Insurance and Finance*, Vol. 33 (Berlin: Springer)
- Enriquez, J. E., Siemion, A., Foster, G., et al. 2017, *ApJ*, **849**, 104
- Fogg, M. J. 1987, *Icar*, **69**, 370
- Forgan, D. H. 2017, *IJAsB*, **16**, 349
- Forgan, D. H. 2019, *Solving Fermi's Paradox* (Cambridge: Cambridge Univ. Press)
- Freitas, R. A. J., & Valdes, F. 1985, *AcAau*, **12**, 1027
- Gajjar, V., Perez, K. I., Siemion, A. P. V., et al. 2021, *AJ*, **162**, 33
- Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, *MNRAS*, **362**, 41
- Garrett, M. A. 2015, *A&A*, **581**, L5
- Garrett, M. A., & Siemion, A. P. V. 2023, *MNRAS*, **519**, 4581
- Geringer-Sameth, A., Koushiappas, S. M., & Walker, M. G. 2015, *PhRvD*, **91**, 083535
- Glade, N., Ballet, P., & Bastien, O. 2012, *IJAsB*, **11**, 103
- Gonzalez, G., Brownlee, D., & Ward, P. 2001, *Icar*, **152**, 185
- Gowanlock, M. G. 2016, *ApJ*, **832**, 38
- Gray, R. H., & Ellingsen, S. 2002, *ApJ*, **578**, 967
- Gray, R. H., & Mooley, K. 2017, *AJ*, **153**, 110
- Griffith, R. L., Wright, J. T., Maldonado, J., et al. 2015, *ApJS*, **217**, 25
- Gulkis, S. 1985, in *IAU Symp. 112, The Search for Extraterrestrial Life: Recent Developments*, ed. M. D. Papagiannis (Dordrecht: Reidel), 411
- Gumbel, E. J. 1958, *Statistics of Extremes* (New York: Columbia Univ. Press),
- Haenggi, M. 2013, *Stochastic Geometry for Wireless Networks* (Cambridge: Cambridge Univ. Press)
- Haigh, J. 2013, *Probability Models* (Berlin: Springer)
- Hair, T. W. 2011, *IJAsB*, **10**, 131
- Hanski, I. 1998, *Natur*, **396**, 41
- Hanson, R., Martin, D., McCarter, C., & Paulson, J. 2021, *ApJ*, **922**, 182
- Haqq-Misra, J. D., & Baum, S. D. 2009, *JBIS*, **62**, 47
- Harp, G. R., Ackermann, R. F., Astorga, A., et al. 2018, *ApJ*, **869**, 66
- Harris, M. J. 1986, *Ap&SS*, **123**, 297
- Hart, M. H. 1975, *QJRAS*, **16**, 128
- Hart, M. H. 1979, *Icar*, **37**, 351
- Harwit, M. 1981, *Cosmic Discovery: The Search, Scope, and Heritage of Astronomy* (New York: Basic Books Inc.)
- Hippke, M. 2018, *AcAau*, **151**, 53
- Hippke, M., & Forgan, D. H. 2017, arXiv:1711.05761
- Hogg, D. W. 1999, arXiv:astro-ph/9905116
- Horowitz, P., & Sagan, C. 1993, *ApJ*, **415**, 218
- Howard, A. W., Horowitz, P., Wilkinson, D. T., et al. 2004, *ApJ*, **613**, 1270
- Huang, S.-S. 1959, *PASP*, **71**, 421
- Imara, N., & Di Stefano, R. 2018, *ApJ*, **859**, 40
- Inoue, M., & Yokoo, H. 2011, *JBIS*, **64**, 59
- Isaacson, H., Siemion, A. P. V., Marcy, G. W., et al. 2017, *PASP*, **129**, 054501
- Johnson, J. A., Aller, K. M., Howard, A. W., & Crepp, J. R. 2010, *PASP*, **122**, 905
- Jones, E. M. 1981, *Icar*, **46**, 328
- Jusup, M., Holme, P., Kanazawa, K., et al. 2022, *PhR*, **948**, 1
- Kardashev, N. S. 1985, in *IAU Symp. 112, The Search for Extraterrestrial Life: Recent Developments*, ed. M. D. Papagiannis (Dordrecht: Reidel), 497
- Karlis, D., & Xekalaki, E. 2005, *International Statistical Review/Revue Internationale de Statistique*, **73**, 35
- Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, *Icar*, **101**, 108
- Kemp, C. D. 1967, *J. Stat. Soc. Inq. Soc. Irel.*, **21**, 151
- Kingman, J. F. C. 1993, *Poisson Processes* (Oxford: Clarendon)
- Kipping, D. 2021, *RNAAS*, **5**, 44
- Kipping, D., Frank, A., & Scharf, C. 2020, *IJAsB*, **19**, 430
- Kipping, D., & Gray, R. 2022, *MNRAS*, **515**, 1122
- Kirsten, F., Marcote, B., Nimmo, K., et al. 2022, *Natur*, **602**, 585
- Klenke, A. 2020, *Probability Theory: A Comprehensive Course* (Cham: Springer)
- Kuiper, T. B. H., & Morris, M. 1977, *Sci*, **196**, 616
- Lacki, B. C. 2016, arXiv:1604.07844
- Lacki, B. C. 2019, *PASP*, **131**, 024102
- Lacki, B. C. 2020, *ApJ*, **905**, 18
- Lacki, B. C. 2021, *IJAsB*, **20**, 359
- Landis, G. A. 1998, *JBIS*, **51**, 163
- Last, G., & Penrose, M. 2017, *Lectures on the Poisson Process* (Cambridge: Cambridge Univ. Press)
- Laughlin, G., Bodenheimer, P., & Adams, F. C. 1997, *ApJ*, **482**, 420
- Lawrence, E., Vander Wiel, S., Law, C., Burke Spolaor, S., & Bower, G. C. 2017, *AJ*, **154**, 117
- Learned, J. G., Kudritzki, R. P., Pakvasa, S., & Zee, A. 2012, *ConPh*, **53**, 113
- Learned, J. G., Pakvasa, S., Simmons, W. A., & Tata, X. 1994, *QJRAS*, **35**, 321
- Leibold, M. A., Holyoak, M., Mouquet, N., et al. 2004, *EcolL*, **7**, 601
- Lin, H. W., Gonzalez Abad, G., & Loeb, A. 2014, *ApJL*, **792**, L7
- Lineweaver, C. H., Fenner, Y., & Gibson, B. K. 2004, *Sci*, **303**, 59
- Lingam, M., Ginsburg, I., & Bialy, S. 2019, *ApJ*, **877**, 62
- Lingam, M., & Loeb, A. 2017, *ApJL*, **837**, L23
- Lingam, M., & Loeb, A. 2020, *ApJ*, **894**, 36
- Lingam, M., & Loeb, A. 2021, *Life in the Cosmos: From Biosignatures to Technosignatures* (Cambridge, MA: Harvard Univ. Press)
- Livio, M. 1999, *ApJ*, **511**, 429
- Maire, J., Wright, S. A., Barrett, C. T., et al. 2019, *AJ*, **158**, 203
- Makovetskii, P. V. 1977, *SvA*, **21**, 251
- Martínez, V. J., & Saar, E. 2002, *Statistics of the Galaxy Distribution* (London: Chapman and Hall)
- Messerschmitt, D. G. 2015, *AcAau*, **107**, 20

- Miller, G. E., & Scalo, J. M. 1979, *ApJS*, **41**, 513
- Moustakas, J., Coil, A. L., Aird, J., et al. 2013, *ApJ*, **767**, 50
- Moyal, J. E. 1962a, *RSPSA*, **266**, 518
- Moyal, J. E. 1962b, *AcMa*, 108, 1
- Mulders, G. D., Pascucci, I., & Apai, D. 2015, *ApJ*, **814**, 130
- Muno, M. P., Pfahl, E., Baganoff, F. K., et al. 2005, *ApJL*, **622**, L113
- Napier, W. M. 2004, *MNRAS*, **348**, 46
- Neyman, J., & Scott, E. L. 1952, *ApJ*, **116**, 144
- Neyman, J., & Scott, E. L. 1958, *J. R. Stat. Soc., Ser. B, Methodol.*, **20**, 1
- Nishino, Y., & Seto, N. 2018, *ApJL*, **862**, L21
- Oliver, B. M., & Billingham, J. 1971, Project Cyclops: A Design Study of a System for Detecting Extraterrestrial Intelligent Life, [NASA-CR-114445NASA](#), NASA
- Olson, S. J. 2015, *CQGra*, **32**, 215025
- Olson, S. J. 2016, *JCAP*, **2016**, 021
- Osmanov, Z. 2016, *IJAsB*, **15**, 127
- Pace, G. W., & Walker, J. C. G. 1975, *Natur*, **254**, 400
- Pooley, D., Lewin, W. H. G., Anderson, S. F., et al. 2003, *ApJL*, **591**, L131
- Prantzos, N. 2020, *MNRAS*, **493**, 3464
- Price, D. C. 2021, in The WSPC Handbook of Astronomical Instrumentation, Volume 1: Radio Astronomical Instrumentation, ed. A. Wolszczan (Singapore: World Scientific), 159
- Price, D. C., Enriquez, J. E., Brzycki, B., et al. 2020, *AJ*, **159**, 86
- Rose, C., & Wright, G. 2004, *Natur*, **431**, 47
- Ross, S. M. 1996, *Stochastic Processes* (New York: Wiley)
- Sagan, C. 1973, *Icar*, **19**, 350
- Sagan, C. 1985, *Contact: A Novel* (New York: Simon & Schuster, Inc.)
- Scheffer, L. K. 1994, *QJRAS*, **35**, 157
- Scheffer, L. K. 2014, *IJAsB*, **13**, 62
- Scheuer, P. A. G. 1957, *PCPS*, **53**, 764
- Schmidt, G. A., & Frank, A. 2019, *IJAsB*, **18**, 142
- Schwartz, R. N., & Townes, C. H. 1961, *Natur*, **190**, 205
- Semiz, I., & Oğur, S. 2015, [arXiv:1503.04376](#)
- Sheikh, S. Z., Smith, S., Price, D. C., et al. 2021, *NatAs*, **5**, 1153
- Sheikh, S. Z., Wright, J. T., Siemion, A., & Enriquez, J. E. 2019, *ApJ*, **884**, 14
- Shields, A. L., Ballard, S., & Johnson, J. A. 2016, *PhR*, **663**, 1
- Shostak, S., Ekers, R., & Vaile, R. 1996, *AJ*, **112**, 164
- Shvartsman, V., Beskin, G., Mitronova, S., et al. 1993, in ASP Conf. Ser. 47, Third Decennial US-USSR Conf. on SETI, ed. G. S. Shostak (San Francisco, CA: ASP), 381
- Siemion, A., Von Korff, J., McMahon, P., et al. 2010, *AcAau*, **67**, 1342
- Smart, J. M. 2009, in *Cosmos & Culture : Cultural Evolution in a Cosmic Context*, ed. S. J. Dick & M. L. Lupisella (Washington, DC: NASA), 201
- Stevenson, D. J. 1999, *Natur*, **400**, 32
- Suazo, M., Zackrisson, E., Wright, J. T., Korn, A. J., & Huston, M. 2022, *MNRAS*, **512**, 2988
- Subotowicz, M. 1979, *AcAau*, **6**, 213
- Tarter, J. 2001, *ARA&A*, **39**, 511
- Tarter, J. C. 1984, *AcAau*, **11**, 387
- Tarter, J. C. 2007, *HiA*, **14**, 14
- Tellis, N. K., & Marcy, G. W. 2017, *AJ*, **153**, 251
- Tipler, F. J. 1980, *QJRAS*, **21**, 267
- Tonry, J., & Schneider, D. P. 1988, *AJ*, **96**, 807
- Tremblay, C. D., & Tingay, S. J. 2020, *PASA*, **37**, e035
- Vidal, C. 2011, [arXiv:1104.4362](#)
- Ward, P., & Brownlee, D. 2000, *Rare Earth: Why Complex Life is Uncommon in the Universe* (New York: Copernicus Books)
- Wasserman, L. 2004, *All of Statistics: A Concise Course in Statistical Inference* (New York: Springer)
- Webb, S. 2015, *If the Universe Is Teeming with Aliens... Where is Everybody?* (Cham: Springer)
- Whitmire, D. P., & Wright, D. P. 1980, *Icar*, **42**, 149
- Wiegand, T., & Moloney, K. A. 2013, *Handbook of Spatial Point-pattern Analysis in Ecology* (London: Chapman and Hall)
- Wiley, K. B. 2011, [arXiv:1111.6131](#)
- Włodarczyk-Sroka, B. S., Garrett, M. A., & Siemion, A. P. V. 2020, *MNRAS*, **498**, 5720
- Worden, S. P., Drew, J., Siemion, A., et al. 2017, *AcAau*, **139**, 98
- Wright, J. T. 2018, *IJAsB*, **17**, 96
- Wright, J. T., Cartier, K. M. S., Zhao, M., Jontof-Hutter, D., & Ford, E. B. 2016, *ApJ*, **816**, 17
- Wright, J. T., Haqq-Misra, J., Frank, A., et al. 2022, *ApJL*, **927**, L30
- Wright, J. T., Kanodia, S., & Lubar, E. 2018, *AJ*, **156**, 260
- Wright, J. T., Mullan, B., Sigurdsson, S., & Povich, M. S. 2014, *ApJ*, **792**, 26
- Zackrisson, E., Calissendorff, P., Asadi, S., & Nyholm, A. 2015, *ApJ*, **810**, 23
- Zackrisson, E., Korn, A. J., Wehrhahn, A., & Reiter, J. 2018, *ApJ*, **862**, 21
- Zwicky, F. 1957, *Morphological Astronomy* (Berlin: Springer)