

Properties of an Entropy-based Signal Receiver with an Application to Ultrasonic Molecular Imaging

M.S. Hughes¹, J.E. McCarthy², J.N. Marsh¹, J.M. Arbeit¹, R.G. Neumann¹, R.W. Fuhrhop¹, K.D. Wallace¹, D.R. Znidarsic¹, B.N. Maurizi², S.L. Baldwin¹, G. M. Lanza¹, and S. A. Wickline¹

¹Washington University School of Medicine

²Department of Mathematics, Washington University

(Dated: February 27, 2007)

Qualitative and quantitative properties of the finite part, H_f , of the Shannon entropy of a continuous waveform $f(t)$ in the continuum limit are derived in order to illuminate its use for waveform characterization. Simple upper and lower bounds on H_f , based on features of $f(t)$, are defined. We also derive quantitative criteria for *a priori* estimation of the average-case variation of H_f and $\log E_f$, where E_f is the signal energy of $f(t)$. These provide relative sensitivity estimates that could be used to prospectively choose optimal imaging strategies in real-time ultrasonic imaging machines, where system bandwidth is often pushed to its limits. To demonstrate the utility of these sensitivity relations for this application, we performed a study designed to assess the feasibility of identification of angiogenic neovasculature targeted with perfluorocarbon nanoparticles that specifically bind to $\alpha_v\beta_3$ -integrin expression in tumors. The outcome of this study agrees with our prospective sensitivity estimates for the two receivers. Moreover, these data demonstrate the ability of entropy-based signal receivers when used in conjunction with targeted nanoparticles to elucidate the presence of $\alpha_v\beta_3$ -integrins in primordial neovasculature, particularly in acoustically unfavorable environments.

I. INTRODUCTION

Detection of molecular epitopes associated with neovasculature in a growing tumor presents a unique challenge for ultrasonic clinical imaging systems. The tumor vasculature is often below the instrument's resolution, which necessitates the use of an enhancing contrast agent. We have demonstrated that a specially modified nanoparticle contrast agent can be successfully targeted towards the expression of $\alpha_v\beta_3$ integrins that are expressed in abundance on tumor neovasculature¹.

The successful detection of cancer *in vivo* depends on various factors when using molecularly targeted contrast agents. The number of epitopes to which the ligand can bind must be sufficient to allow enough of the contrast agent to accumulate for detection, and ligand specificity must be maintained to ensure that nonspecific binding remains negligible. In addition, the pharmacokinetics of the contrast agent must be such that the agent remains in circulation long enough to bind in sufficient quantity to the desired target. Ideally, the background signal from unbound, circulating contrast agent is low enough or absent so as to not interfere with the assessment of bound, targeted agent. The imaging technology itself also must be highly sensitive and capable of detecting and/or quantifying the level of contrast agent bound to the pathological tissue. In clinical ultrasonic imaging, the sensitivity of detection depends on a physical difference in the way sound interacts with a surface covered by targeted contrast agent versus one that is not. The data presented below show that in many cases, the sensitivity of this determination can be improved by applying novel and specific signal processing techniques based on thermodynamic or information-theoretic analogs.

Previous work already has demonstrated for high frequency ultrasound in epitope-rich pathologies, such

as fibrin in thrombus, that targeted perfluorocarbon nanoparticles act as a suitable molecular imaging agent by modifying the acoustic impedance on the surface to which they bind². However, at lower frequencies and for sparse molecular epitopes, clear delineation between non-targeted normal tissue and angiogenic vessels remains a challenge.

Site-targeted nanoparticle contrast agents, when bound to the appropriate receptor, must be detected in the presence of bright echoes returned from the surrounding tissue. In this study, we attempt to solve the problem of detection of site-specific contrast through the use of signal receivers (*i.e.*, mathematical operations that reduce an entire radio frequency (RF) waveform or a portion of it to a single number) based on information-theoretic quantities, such as Shannon Entropy (H), or its counterpart for continuous signal (H_f). These receivers appear to be sensitive to diffuse, low amplitude features of the signal that often are obscured by noise, or else lost in large specular echoes and, hence, not usually perceivable by a human observer.³⁻⁷

Although entropy-based techniques have a long history in image processing for image enhancement and post processing of reconstructed images, the approach we take is different in that entropy is used directly as the quantity defining the pixel values in the image. Specifically, images are reconstructed by computing the entropy (or a limiting form of it: H_f) of segments of the individual RF A-lines that comprise a typical medical image by applying a "moving window", or "box-car", analysis. The computation of an entropy value for each location within an image is therefore possible, and the results can be superimposed over the conventional grayscale image as a parametric map.

For comparison, we also used more conventional signal processing techniques on the same beam formed RF:

specifically, the signal energy and its logarithm. This was done after a preliminary B-mode grayscale analysis of backscattered data acquired for this study was unable to detect changes in backscattered RF arising from the accumulation of targeted nanoparticles in the tumor neovasculature. This result implied that acoustic characterization of sparse collections of targeted perfluorocarbon nanoparticles presented challenges that might require the application of novel types of signal processing.

II. APPROACH

All RF data are obtained by sampling a continuous function $y = f(t)$, which has associated with it two functions: its density function, $w_f(y)$, and its increasing rearrangement, $f_{\uparrow}(t)$. Either may be used to compute H_f and to provide bounds for the relative magnitudes of changes of H_f with respect to familiar measures such as energy, E_f , and log energy, $\log(E_f)$.

A. The function $w_f(y)$

The density function $w_f(y)$ may be used to compute the entropy H_f and the signal energy E_f (and hence also $\log[E_f]$). It corresponds to the density functions that are the primary mathematical object in much of statistical signal processing and from which other mathematical quantities are subsequently derived (*e.g.*, mean values, variances, covariances)^{8–10}. In that setting, it constitutes the most fundamental unit of information that the experimentalist has about a random variable, which is often implicitly assumed to have a form that is too complicated to be worth analyzing (hence the density function is analyzed instead). Moreover, this function is usually assumed to be very well-behaved mathematically, that is, to be continuous, infinitely differentiable, and to approach zero at infinity. On the other hand, in our analysis $f(t)$ is the fundamental unit of information and $w_f(y)$ is a derived quantity obtained from it. As we shall see, $w_f(y)$ is not particularly well-behaved mathematically. In spite of the difficulties this adds to the mathematical analysis, several experimental studies indicate that signal processing using the derived $w_f(y)$ permits definition of signal receivers that are more sensitive than energy-based receivers. We will also show that signal energy can be computed using $w_f(y)$, and thus all of conventional energy-based signal processing may be placed within the mathematical framework described below.

Without loss of generality, we may adopt the convention that the domain of $f(t)$ is $[0, 1]$. Then, $w_f(y)$, the density function of $f(t)$, can be defined by the basic integral relation

$$\int_0^1 \phi(f(t)) dt = \int_{f_{min}}^{f_{max}} \phi(y) w_f(y) dy, \quad (1)$$

for any continuous function $\phi(y)$. This should be compared with the expression for the expectation value of a function ϕ of a random variable X with density $p_X(x)$, which is given by

$$\int \phi(x) p_X(x) dx,$$

which explains why $w_f(y)$ is referred to as the density function for $f(t)$.¹¹ Most applications of either probability or information theory to signal processing proceed, usually very early in the discussion, by assuming a specific density function^{12–14}. In contrast, the analysis described in this study begins with a digitized time-domain function and derives its density function.

We will see that different functions may have the same density function; these are referred to as “equidistributed” functions.

If $\phi(y) = 1$, Eq. (1) implies,

$$\int_{f_{min}}^{f_{max}} w_f(y) dy = 1, \quad (2)$$

a relation that will be used below. If we set $\phi(y) = \delta(y - \xi)$ then

$$\int_0^1 \delta(f(t) - \xi) dt = \int_{f_{min}}^{f_{max}} \delta(y - \xi) w_f(y) dy = w_f(\xi). \quad (3)$$

Thus, using the well-known expression

$$\int_0^1 \delta(f(t) - \xi) dt = \sum_{\{t_k | f(t_k) = \xi\}} \frac{1}{|f'(t_k)|}, \quad (4)$$

and Eqs. (3, 4) we obtain

$$w_f(\xi) = \sum_{\{t_k | f(t_k) = \xi\}} \frac{1}{|f'(t_k)|}, \quad (5)$$

an expression for w_f that reveals some of its qualitative features in terms of $f(t)$. For instance, $w_f(y) \geq 0$ for all y ; it becomes $w_f(y - c)$ if we add a constant, c , to $f(t)$; $w_{-f}(-y) = w_f(y)$; and perhaps most importantly $w_f(y)$ has singularities at values of $y = f(t)$ where $f'(t) = 0$. We will assume that all digitizable waveforms $f(t)$ are comprised of at least one section, or “lap”, where it is monotonic. The lap boundaries are just the points t where $f'(t) = 0$. On each lap, $f(t)$ has a well-defined inverse function. These may be used to rewrite Eq. (5)

$$w_f(y) = \sum_{k=1}^N |g'_k(y)|, \quad (6)$$

where N is the number of laps, $g_k(y)$ is the inverse of $f(t)$ in the k^{th} -lap and if y is not in the range of $f(t)$ in the k^{th} -lap, $g'_k(y)$ is taken to be 0. This equation may also be rewritten in a more explicit form as

$$w_f(y) = \sum_{k=1}^N |g'_k(y)| \Pi(y_{L_i}, y_{U_i}, y), \quad (7)$$

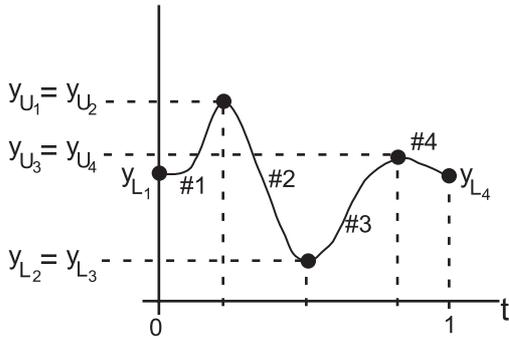


FIG. 1. An illustration of the y_{L_i} and y_{U_i} in Eq. (7) for a four lap function. $y_{L_1} = f(0)$, $y_{L_1} = y_{L_2}$, $y_{U_2} = y_{U_3}$, $y_{L_3} = y_{L_4}$, and $y_{U_4} = f(1)$.

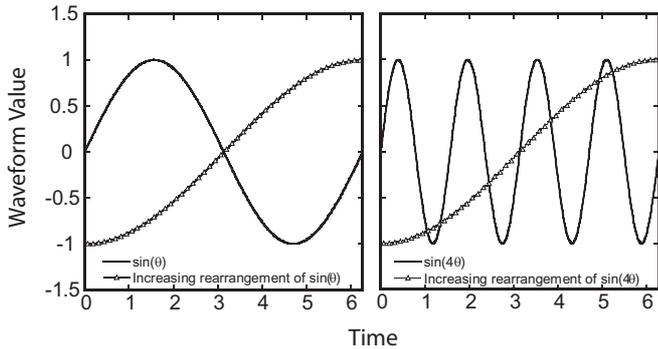


FIG. 2. Rearrangements of $\sin(2\pi t)$ and $\sin(8\pi t)$. Note that the increasing rearrangements are the same as we would expect since $\sin(2\pi t)$ and $\sin(8\pi t)$ are equimeasurable.

where y_{L_i} , y_{U_i} denote the values of $f(t)$ at the lower and upper bounds of the i^{th} lap respectively, and $\Pi(y_{L_i}, y_{U_i}, y)$ denotes a Heaviside-like step-function that is one for $y \in [y_{L_i}, y_{U_i}]$ and zero otherwise. These conventions are illustrated in Figure (1). This form will prove useful below.

Application of Eq. (6) to the computation of the density functions of $\sin(n\pi t)$, where n is an integer, reveals that they are all equal. Figure (2), shows two such functions and approximations to their increasing rearrangements, $f_{\uparrow}(t)$, obtained by sorting sampled (at 100 points) versions of both functions. As shown in the figure, waveforms with drastically different frequency content may be equidistributed and have equal entropy, H_f . Thus, frequency plays little, if any, direct role in our analysis.

For experimentally measured data, we may also assume that all digitizable waveforms $f(t)$ have a Taylor series expansion at all points in $[0, 1]$. Then near a time t_k such that $f'(t_k) = 0$

$$y = f(t) = f(t_k) + \frac{1}{2!} f''(t_k) (t - t_k)^2 + \dots, \quad (8)$$

t_k is a lap boundary and on the left side of this point Eq. (8) may be truncated to second order and inverted

to obtain

$$g_k(y) = \pm \sqrt{2(y - f(t_k)) / f''(t_k)}, \quad (9)$$

with

$$|g'_k(y)| = 1 / \sqrt{2f''(t_k)(y - f(t_k))}. \quad (10)$$

The contribution to $w_f(y)$ from the right side of the lap boundary, from $g_{k+1}(y)$, is the same, so that the overall contribution to $w_f(y)$ coming from the time interval around t_k is

$$|g'(y)| = \sqrt{2 / (f''(t_k)(y - f(t_k)))}, \quad (11)$$

and so $w_f(y)$ has only a square root singularity (we have assumed that t_k is interior to the interval $[0, 1]$, if not, then the contributions to w_f come from only the left or the right). If additionally, $f''(t_k) = 0$ then the square root singularity in Eq. (10) will become a cube-root singularity, and so on, so that the density functions we consider will have only integrable algebraic singularities.

Equations (6), (7) and (10), (11) determine the qualitative features of a typical density function, $w_f(y)$. From Eqs. (6) and (7) we see that the density function may exhibit finite jump discontinuities. Equations (10) and (11) imply further, that algebraic singularities may also be present.

Figure (3) illustrates, schematically, several possible types of behaviors possible in $w_f(y)$: both discontinuities and algebraic singularities (indicated by arrows on the plots of $w_f(y)$). Progressing from left to right in any given row also illustrates how to estimate qualitative features of $w_f(y)$ from $f(t)$. For instance, in the top row the maxima in $f(t)$ correspond to an algebraic singularity in $w_f(y)$, plotted sideways in the middle column to more clearly indicate the relationship between its features and those of $f(t)$. The third column shows $w_f(y)$ in a conventional layout (a rotated and flipped version of the plot in the middle column). These plots show that the density functions possess significantly different attributes from those usually considered in statistical signal processing.

The mathematical characteristics of the singularities are important in order to guarantee the existence of the following integral on which we base our analysis of signals:

$$H_f = \int_{f_{\min}}^{f_{\max}} w_f(y) \log w_f(y) dy. \quad (12)$$

In previous studies, this quantity was referred to as H_C , since it is obtained from the Shannon entropy, H , in the limit where a waveform, $f(t)$ is acquired with infinite sampling rate and dynamic range (*i.e.*, in a limit approaching the continuum).³⁻⁷ However, in the current study, which emphasizes the changes in entropy as the underlying waveform $f(t)$ changes, it seems more appropriate to use the notation H_f to emphasize its dependence on $f(t)$.

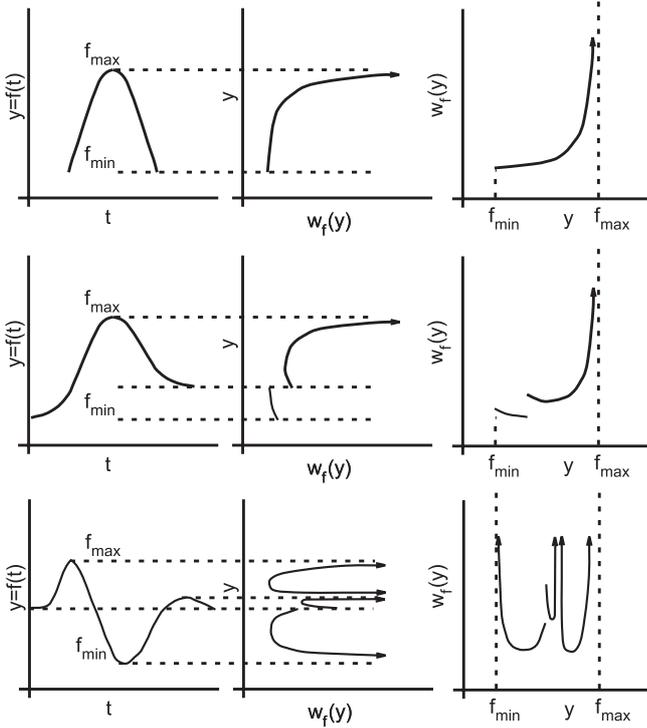


FIG. 3. Plots of three different time-domain waveforms, $f(t)$ (left column), and their associated density functions $w_f(y)$ (middle column and right columns).

Typically, ultrasonic signal analysis has been based on analysis of backscattered energy or its logarithm.^{3,5,7} For instance, the gray-scale images used in medical ultrasound B-mode displays of the logarithm of the envelope (*i.e.*, analytic signal magnitude) of the backscattered signal, which is well-known to represent the energy contained in the waveform over the sampling interval.¹⁵ Energy contained in longer intervals may be estimated using the sum of the squares of digitized values contained in a “moving window”. The physical motivation is based on analogy with the kinetic energy of a spring, which is proportional to amplitude of spring extension squared. This quantity, E_f , may also be expressed using $w_f(y)$ via Eq. (1)

$$E_f = \int_0^1 f(t)^2 dt = \int_{f_{min}}^{f_{max}} w_f(y) y^2 dy. \quad (13)$$

B. The function $f_{\uparrow}(t)$

Like signal energy, E_f , the entropy may also be computed directly from a time domain function, $f_{\uparrow}(t)$ (f “increasing”). In addition, approximations of this function may be found quickly by sorting discretely sampled data from $f(t)$. This function may also be used to rapidly

estimate the relative sensitivities of entropy and energy-based receivers (which we will derive in Eq. (56) and Eq. (57)) and is thus useful for qualitative evaluation of entropy receiver performance.

Precise definition of $f_{\uparrow}(t)$ begins with the observation that since $w_f(y) \geq 0$, the function defined by the integral

$$W_f(y) = \int_{f_{min}}^y w_f(s) ds \quad (14)$$

is nondecreasing. Its inverse, a function of time denoted $f_{\uparrow}(t)$, is equimeasurable with $f(t)$ by Eq. (5) and the inverse function theorem. Thus, by Eq. (5), $1/f'_{\uparrow}(t) = w_f(f_{\uparrow}(t))$, so that

$$H_f = - \int_0^1 \log [f'_{\uparrow}(t)] dt. \quad (15)$$

The function $f_{\uparrow}(t)$ is referred to as the increasing rearrangement of $f(t)$. It may be approximated quickly, and with reasonable accuracy, by sorting the digitized values of $f(t)$.

Since the function $f_{\uparrow}(t)$ is equimeasurable with $f(t)$ it follows that $f_{\uparrow max} = f_{max}$ and $f_{\uparrow min} = f_{min}$. The increasing rearrangement has been useful for rapid computation of H_f . The calculation uses the Green’s function for the second order differential operator to simultaneously smooth and differentiate twice the approximation to $f_{\uparrow}(t)$ obtained by sorting the sampled values of $f(t)$.⁷

C. Physical interpretation of $w_f(y)$ and its numerical calculation

One challenge of this analysis is to achieve an intuitive appreciation for the physical meaning of entropy that relates to imaging applications and the detection of subtle pathologies. There are at least two different paths to obtaining a physical meaning for the quantity H_f appearing in Eq. (12) (or Eq. (15)). One is based on the fact that H_f is derived from the Shannon entropy, which has been extensively investigated in communication theory. While this connects H_f to the capacity of a communications channel, which is a quasi-physical quantity, it also implies a relation with Kolmogorov complexity, an uncomputable quantity that is often approximated by the Shannon entropy. The physical significance of Kolmogorov complexity and its use in deriving Lagrangians for both classical mechanics and special relativity has been described in a recent article by Soklakov.¹⁶

The second approach is to understand the physical meaning of the density function $w_f(y)$ which we now discuss. Figure (4) is a diagram of a very idealized ultrasonic backscatter measurement using a single point-source transducer, which we assume to have infinite bandwidth so that it emits δ -function pulses. We suppose that the medium is composed of particulate scatterers that backscatter like perfect reflectors with either no phase inversion (*e.g.*, fixed boundary) or 180° phase

inversion (*e.g.*, free boundary). The result is that the backscattered ultrasound will be comprised of a train of sums of positive and negative delta functions. The Figure inset shows the output at one instant that is the sum of two positive (A and C) and one negative (B) δ -function. These are the reflections of the original transmit pulse from the scatterers on the spherical shell of constant time centered on the transducer face. The backscattered signal will, in the idealized case considered here, be a train of these pulses, with the height of each pulse representing the sign-weighted sum of particle scatterers on each constant-time spherical shell. The received signal from the transducer is a train of such pulses, with the received amplitude at two instants being the same when the sign-weighted sum in the corresponding shells is equal. The number of times that the output amplitude is the same is the density $w_f(y)$ for this idealized δ -function pulse train waveform. In reality, the transducer has finite bandwidth, the reflectors are neither perfect nor discrete and the measured waveform is a convolution of transducer and scatterer transfer functions. Nevertheless, the idealized function approach sheds light on the physical interpretation of $w_f(y)$, as an approximate representation of the spatial distribution of scatterers in the acoustic field.

D. Intrinsic vs. extrinsic statistics

While our approach may have certain similarities to conventional statistical signal processing, it also differs significantly in several ways as will be illustrated in the next few sections. These differences fall into two categories: mathematical and physical.

Unlike most statistical signal processing, where the underlying density is the starting point of the analysis, our analysis begins in the time domain, with a function like $\sin(t)$, $\sin(kt)$ or a weighted sum of these functions and derives their density function. Second, this is done for each waveform and not, as is typically the case (*e.g.* in derivation of the Rayleigh or Rician distributions), for an ensemble of waveforms. Moreover, the density functions we consider contain algebraic singularities and hence cannot typically be convolved. This eliminates the possibility of applying the central limit theorem in our analysis, which contrasts sharply with traditional statistical signal analysis, where density functions are bounded and the central limit theorem often plays a significant role. Third, the waveforms we consider are not typically independent functions, for instance $\sin(t)$ and e^{-t^2} are not independent¹⁷⁻¹⁹. Thus, we may not compute the density of a sum of waveforms from the convolution of their density functions. There are three reasons for adopting the approach described in this study. First, the assumption of statistical independence can lead to conclusions that are in some ways at “odds” with the usual assumptions of eigenfunction-based signal processing. Second, independent functions are “rare” in the sense discussed in Appendix A.2. Third, as shown in the results section

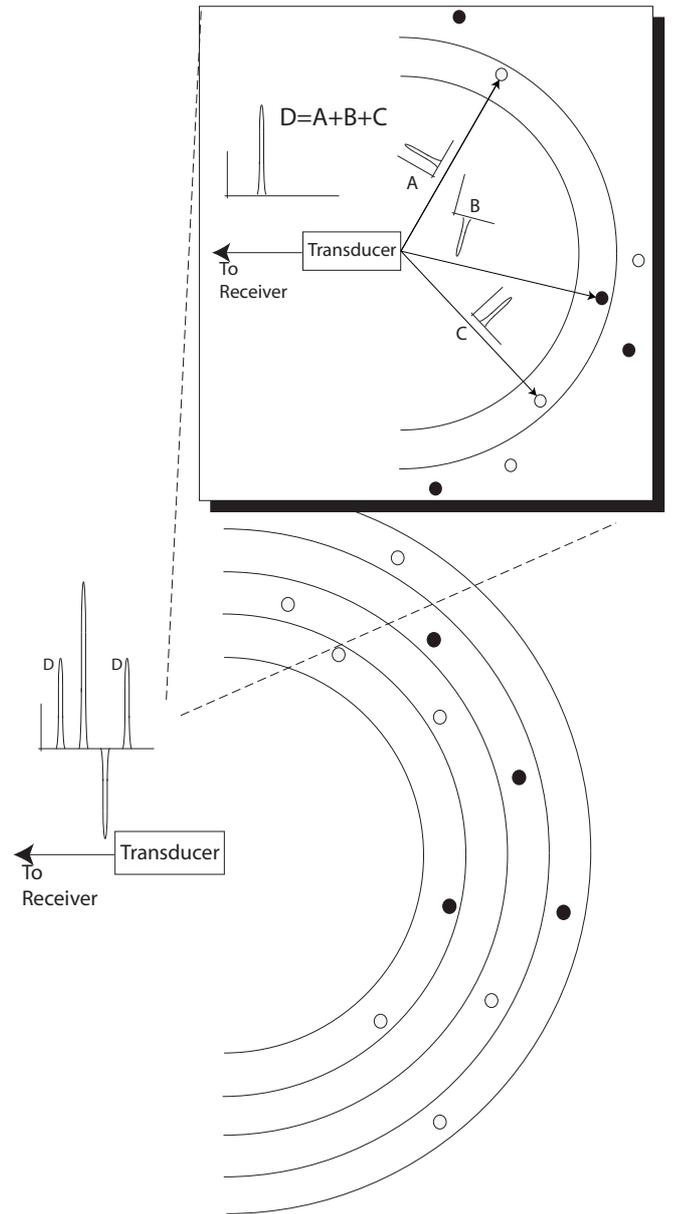


FIG. 4. An idealized scattering experiment using a transducer emitting perfect δ -function pulses. This model illuminates the physical interpretation of the function $w_f(y)$ used to compute H_f .

and in several other studies, this approach can lead to significant improvements in detection of subtle changes in backscattered ultrasonic waveforms^{3-7,20-24}.

Most applications of statistics to signal processing follow one of two paths. The simplest is to assume that a canonical noise term, *e.g.*, Gaussian or Poisson, has been added to the signal^{8-10,25-39}. The more physical analysis derives a functional form for the statistical density of the waveform, (*e.g.*, Rician, Rayleigh, K-distribution) based on an idealized “thought” experiment consisting of many different random scattering configurations⁴⁰. As

the number of these virtual experiments grows to infinity a well defined density is obtained. Waveforms are then characterized externally, by reference to this density. In the approach described here, the density of values of individual waveforms assumes central importance. Each waveform is characterized intrinsically, using its internal statistics as captured by the density function $w_f(y)$. Our motivation is the observation that a shift from extrinsic to intrinsic viewpoint, has proven useful in the past: algorithmic information theory (*i.e.*, intrinsic viewpoint) was developed in the 1960's, following statistical information theory (*i.e.*, extrinsic viewpoint).

To pursue this approach a large number of the simplifying assumptions typically applied in statistical signal processing must be abandoned: independence, normality, continuity of density functions. This increases the mathematical complexity of our analysis. However, much of traditional signal processing, in fact all of it that is based on integrals of the digitized waveforms, may be placed in the context of this formalism, as shown by Eq. (13). Moreover, as we will show below, significant gains in experimental sensitivity may be gained that make the extra effort worthwhile.

E. Scaling invariance of ΔH_f and $\Delta \log E_f$

The scaling properties of a signal receiver are also important since in a typical ultrasound experiment gain settings on imager or apparatus amplifiers are often adjusted to optimize dataset dynamic range. As we will see, both ΔH_f and $\Delta \log E_f$ are invariant under scale changes. This makes their interpretation in experimental studies straightforward. To obtain this result we will first obtain the scaling behavior of $w_f(y)$. From Eq. (5) we see that if we scale $f(t)$ by a factor of σ

$$w_{\sigma f}(\xi) = \sum_{\{t_k | \sigma f(t_k) = \xi\}} \frac{1}{|\sigma| |f'(t_k)|} = \frac{1}{|\sigma|} w_f\left(\frac{\xi}{\sigma}\right), \quad (16)$$

or

$$|\sigma| w_{\sigma f}(\sigma y) = w_f(y), \quad (17)$$

which we may apply in

$$\begin{aligned} H_f &= \int_{f_{min}}^{f_{max}} w_f(y) \log [w_f(y)] dy \\ &= \int_{f_{min}}^{f_{max}} |\sigma| w_{\sigma f}(y) \log [|\sigma| w_{\sigma f}(y)] dy \\ &= \int_{\sigma f_{min}}^{\sigma f_{max}} w_{\sigma f}(\xi) \log [|\sigma| w_{\sigma f}(\xi)] d\xi \\ &= \int_{\sigma f_{min}}^{\sigma f_{max}} w_{\sigma f}(\xi) \log [w_{\sigma f}(\xi)] d\xi + \log [|\sigma|] \\ &= H_{\sigma f} + \log [|\sigma|], \end{aligned} \quad (18)$$

which tells us that if we perform a measurement where $f(t) \rightarrow \tilde{f}(t)$ to obtain ΔH_f , and then repeat the measurement at a different gain setting to measure $\sigma f(t) \rightarrow \sigma \tilde{f}(t)$ we will obtain the same change in entropy *i.e.*,

$$\Delta H_{\sigma f} = \Delta H_f. \quad (19)$$

The same is true for $\log[E_f]$ since

$$\begin{aligned} \log E_{\sigma f} &= \int_0^1 \log [\sigma f(t)]^2 dt \\ &= 2 \log |\sigma| + \log E_f \end{aligned} \quad (20)$$

so that

$$\Delta \log [E_{\sigma f}] = \Delta \log [E_f], \quad (21)$$

which states that signal energy is the second moment of $w_f(y)$.

F. Qualitative features of H_f

1. Inequalities bounding H_f from below

There are two inequalities useful for qualitative estimation of H_f from $f(t)$. The first of these is based on Gibbs' inequality,

$$\int_0^1 w_f(y) \log h(y) dy \leq \int_0^1 w_f(y) \log w_f(y) dy = H_f, \quad (22)$$

where the functions $w_f(y)$ and $h(y)$ satisfy

$$\int_0^1 h(y) dy = \int_0^1 w_f(y) dy = 1. \quad (23)$$

If we let $h(y) = e^{g(y)} / \int_0^1 e^{g(s)} ds$, then

$$\int_0^1 w_f(y) g(y) dy - \int_0^1 w_f(y) dy \int_0^1 e^{g(s)} ds \leq H_f, \quad (24)$$

which for the particular choice $g(y) = y^2$, becomes

$$\begin{aligned} \int_0^1 f^2(t) dt - .38025 &\leq H_f, \\ E - .38025 &\leq H_f, \end{aligned} \quad (25)$$

where E is the signal energy.

We also have

$$\begin{aligned} V[f] &= \int_0^1 |f'_\uparrow(t)| dt = \int_0^1 \sqrt{f'_\uparrow(t)} dt \\ &\leq \int_0^1 \log [f'_\uparrow(t)] dt = -H_f \end{aligned} \quad (26)$$

implying that

$$H_f \geq -V[f] \quad (27)$$

where $V[f]$ is the variation of $f(t)$. Somewhat more interesting is a bound based on the length of $f(t)$ obtained by considering the lengths of all functions $g(t)$ that are equidistributed with $f(t)$, which may be written as (using the notation $f(t) \sim g(t)$ to indicate that f and g are equidistributed):

$$\begin{aligned} \min_{\{g(t)|g(t)\sim f(t)\}} \int_0^1 \sqrt{g(t)^2 + 1} dt &= \int_0^1 \sqrt{f'_\uparrow(t)^2 + 1} dt \\ &\geq \int_0^1 \log f'_\uparrow(t) dt \geq -H_f \end{aligned} \quad (28)$$

implying that

$$H_f \geq -L[f(t)] \quad (29)$$

where $L[f(t)]$ is the length of the graph of $f(t)$ over the interval $[0, 1]$.

2. Inequality bounding H_f from above

We also have

$$\begin{aligned} -H_f &= \int_0^1 \log[f'_\uparrow(t)] dt \\ &\geq \log \int_0^1 f'_\uparrow(t) dt = \log [f_\uparrow(1)] \end{aligned} \quad (30)$$

or

$$H_f \leq \log \left[\frac{1}{f_\uparrow(1)} \right] \quad (31)$$

3. An inequality between ΔH_f and $\Delta \log E_f$

In this section we will derive conditions for the variation of the entropy, ΔH_f , to be greater than that of the log energy, $\Delta \log E_f$. The estimate derived in this section rests on the assumption that $H_f < 0$. While H_f can take on both positive and negative values for the data considered in this study, its value tends to be smallest for sub-segments of backscattered rf corresponding to anatomical regions where accumulation of targeted nanoparticles is expected (as verified by histological staining *e.g.*, Figure (8)). For this reason the inequality is useful for *a priori* estimation of the sensitivities of the two receivers. These can be useful in real-time imaging applications where system bandwidth is often pushed to its limits.

If the waveform, $f(t)$, is perturbed slightly (e.g. by the slow accumulation of nanoparticles in the insonified region), then its density function $w_f(y)$ will also change to $w_f(y) + \varepsilon h(y)$, where we know that

$$\int_{f_{min}}^{f_{max}} w_f(y) dy = \int_{f_{min}}^{f_{max}} w_f(y) + \varepsilon h(y) dy = 1 \quad (32)$$

so that

$$\int_{f_{min}}^{f_{max}} h(y) dy = 0 \quad (33)$$

The signal energy E and the entropy H_f will therefore also change.

$$H_f = \int_{f_{min}}^{f_{max}} w_f(y) \log w_f(y) dy$$

becomes

$$\tilde{H}_f = \int_{f_{min}}^{f_{max}} [w_f(y) + \varepsilon h(y)] \log [w_f(y) + \varepsilon h(y)] dy \quad (34)$$

and

$$E_f = \int_{f_{min}}^{f_{max}} y^2 w_f(y) dy$$

becomes

$$\tilde{E}_f = \int_{f_{min}}^{f_{max}} y^2 [w_f(y) + \varepsilon h(y)] dy. \quad (35)$$

The derivatives of these expressions with respect to $\varepsilon \rightarrow 0$:

$$\begin{aligned} \frac{dH_f}{d\varepsilon} &= \int_{f_{min}}^{f_{max}} h(y) dy + \int_{f_{min}}^{f_{max}} h(y) \log [w_f(y)] dy \\ &= \int_{f_{min}}^{f_{max}} h(y) \log [w_f(y)] dy, \end{aligned} \quad (36)$$

$$\frac{dE_f}{d\varepsilon} = \int_{f_{min}}^{f_{max}} y^2 h(y) dy, \quad (37)$$

may be used to estimate the change in magnitude of the entropy and energy

$$|\Delta H_f| = \varepsilon \left| \frac{dH_f}{d\varepsilon} \right| = \left| \int_{f_{min}}^{f_{max}} h(y) \log [w_f(y)] dy \right|, \quad (38)$$

$$|\Delta E_f| = \varepsilon \left| \frac{dE_f}{d\varepsilon} \right| = \left| \int_{f_{min}}^{f_{max}} y^2 h(y) dy \right|, \quad (39)$$

in the limit where $\varepsilon \rightarrow 0$. Given the perturbation $h(y)$, we wish to obtain an estimate of the relative magnitudes of the integrals appearing in Eqs. (38) and (39) and to establish conditions on $w_f(y)$ under which $|\Delta H_f| \geq |\Delta E_f|$ and under which $|\Delta H_f| \geq |\Delta \log [E_f]|$. *A priori*, we know that the inequality can not hold in general since for $f(t) = t$ we have $w_f(y) = 1$ in which case $|\Delta H_f| = 0$. However, we may still obtain an estimate of average behavior by exploiting the geometric nature of Eqs. (38) and (39) to rewrite them as:

$$|\Delta H_f| = \langle \log^2 [w_f(y)], h(y) \rangle \quad (40)$$

$$= \|\log^2 [w_f(y)]\|_2 \|h(y)\|_2 |\cos(\theta)| \quad (41)$$

$$|\Delta E_f| = \langle y^2, h(y) \rangle = \|y^2\|_2 \|h(y)\|_2 |\cos(\phi)|. \quad (42)$$

where $\langle \bullet, \bullet \rangle$, and $\|\bullet\|_2$ denote the inner product and L_2 norm respectively in the infinite dimensional vector space (*i.e.*, Hilbert space) containing the vectors $h(y)$, $\log^2[w_f(y)]$, and y^2 having magnitudes $\|\log^2[w_f(y)]\|_2$, $\|h(y)\|_2$, $\|y^2\|_2$, and subtending angles θ and ϕ . We suppose that $\log^2[w_f(y)]$ is fixed and that $h(y)$ may vary arbitrarily over the phase space and ask what the average values of the inner product appearing in Eqs. (41) and (42) will be. Geometrically, each average is obtained by sweeping the “random” $h(y)$ vector over all angles subtending the infinite dimensional unit sphere. The averages appearing in Eqs. (41) and (42) then become

$$|\Delta H_f|_{avg} = \|\log^2[w_f(y)]\|_2 \|h(y)\|_2 |\cos(\theta)|_{avg} \quad (43)$$

$$|\Delta E_f|_{avg} = \|y^2\|_2 \|h(y)\|_2 |\cos(\phi)|_{avg}. \quad (44)$$

If we assume that the perturbation $\|h(y)\|_2$ is fixed, then since the average values: $|\cos(\theta)|_{avg}$, $|\cos(\phi)|_{avg}$ over the hypersphere are equal, we see that the relationship between $|\Delta H_f|_{avg}$ and $|\Delta E_f|_{avg}$ is governed by the relation between the quantities,

$$\|\log^2[w_f(y)]\|_2 = \left[\int_{f_{min}}^{f_{max}} [\log(w_f(y))]^2 dy \right]^{1/2}, \quad (45)$$

$$(46)$$

and

$$\begin{aligned} \|y^2\|_2 &= \left[\int_{f_{min}}^{f_{max}} (y^2)^2 dy \right]^{1/2} \\ &= \left[\frac{f_{max}^5 - f_{min}^5}{5} \right]^{1/2}. \end{aligned} \quad (47)$$

Suppose that $H_f < 0$ then

$$\begin{aligned} |H_f| &\leq \left| \int_{w_f(y) \leq 1} w_f(y) \log(w_f(y)) dy \right| \\ &\leq \left(\int_{w_f(y) \leq 1} (w_f(y))^2 dy \right)^{1/2} \\ &\quad \times \left(\int_{w_f(y) \leq 1} (\log(w_f(y)))^2 dy \right)^{1/2} \\ &\leq \left(\int_{w_f(y) \leq 1} w_f(y) dy \right)^{1/2} \\ &\quad \times \left(\int_{w_f(y) \leq 1} (\log(w_f(y)))^2 dy \right)^{1/2} \\ &\leq 1 \left(\int_{w_f(y) \leq 1} (\log(w_f(y)))^2 dy \right)^{1/2} \\ &\leq \left(\int_{f_{min}}^{f_{max}} (\log(w_f(y)))^2 dy \right)^{1/2} \\ &= \|\log[w_f(y)]\|_2 \end{aligned} \quad (48)$$

thus, if $H_f < 0$ and $|H_f|^2 > (f_{max}^5 - f_{min}^5)/5$, *i.e.*, $H_f < -(f_{max}^5 - f_{min}^5)/\sqrt{5}$, then on average $|\Delta H_f| > |\Delta E_f|$. If moreover, $E_f > 1$, so that $\log(E_f) > 0$, then

$$|\Delta H_f| > |\Delta E_f| > |\Delta E_f|/E_f = |\Delta \log E_f|. \quad (49)$$

This average-case relationship is useful for estimation of relative sensitivities of entropy and energy-based receivers as we will show in Figure (13).

4. Rapid Estimation of Relative Sensitivity from $f_{\uparrow}(t)$ and $w_f(y)$

There are several ways to rapidly estimate the relative sensitivities of H_f and $\log E$ without executing the complete calculation of entropy. These are based on estimation of the minimum of $|\Delta H_f|$, (which is relatively expensive to compute directly) which is then compared to the maximum of $|\Delta \log[E_f]|$ (which is easy to compute and will not be discussed further) to obtain a “worst-case” comparison of receiver performances. If $|\Delta \log[E_f]| > |\Delta H_f|$, these estimates may be used to avoid the relatively expensive computation of H_f .

a. Relative Sensitivity from the minimum of $w_f(y)$

An obvious estimate of $|\Delta H_f|$ from the integral appearing in Eq. (45) may be found by computing

$$|\Delta H_f| \approx \sqrt{\min\{\log^2[w_f(y)]\} \times (f_{max} - f_{min})}. \quad (50)$$

Some care must be taken in application of this approximation, since the minimum may occur near a discontinuity. In this case, the Fourier series for w_f will “ring”, which can further reduce the accuracy of Eq. (50), causing it to either over or under estimate $|\Delta H_f|$. This effect can be reduced in the obvious way by increasing the length of Fourier series used to approximate $w_f(y)$, although at the expense of greater computational cost. A further, cruder estimate of $|\Delta H_f|$ may be based upon Eq. (50) by replacing $\min\{\log^2[w_f(y)]\}$ by $\min\{\log^2[f'(t)]\}$ permitting estimation of H_f sensitivity from

$$|\Delta H_f| \approx \min\{\log^2[f'(t)]\} \times (f_{max} - f_{min}). \quad (51)$$

This estimate has the advantages that it is not subject to “ringing” effects and that it may be calculated directly from $f(t)$.

b. Relative Sensitivity from the maximum of $f'_{\uparrow}(t)$

A more sophisticated estimate may be calculated directly from the increasing rearrangement of $f(t)$, $f_{\uparrow}(t)$ and its derivatives, which we write as

$$g(t) = \frac{d}{dt}[f_{\uparrow}(t)], \quad (52)$$

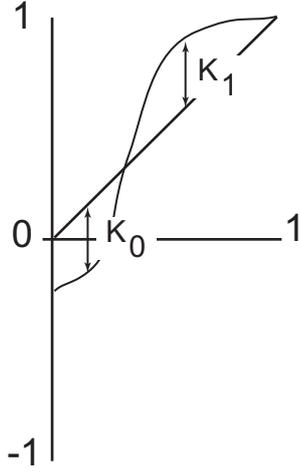


FIG. 5. Definitions of K_0 , K_1 and t_0 , t_1 used to derive the lower bounds on $\|\log[w_f(y)]^2\|_2$ appearing in Eqs. (56) and (57).

and from which we obtain the following expression for the magnitude of H_f variation appearing in Eq. (48)

$$\int_0^1 g(t) \log^2[g(t)] dt = \int_{f_{min}}^{f_{max}} \log^2[w_f(y)] dy. \quad (53)$$

c. Relative Sensitivity from the $f(t)$: K_0 , K_1

As is shown in Appendix B, smoothing of $g(t)$ reduces the magnitude of the integral in Eq. (48). We apply this fact by “smoothing” $g(t)$ to two different constant levels. This may be done by using the maximum distance, K_0 of $g(t)$ below the line $y = t$ (or by using the maximum distance, K_1 of $g(t)$ above the line $y = t$), as defined in Figure (5). The time at which this maximum occurs, t_0 (or t_1) is used to break the interval into two pieces. We replace $g(t)$ by its average value over these two intervals to obtain a “smoothed” function

$$g_1(t) = \begin{cases} \frac{1}{t_0} \int_0^{t_0} g(t) dt = \frac{[t_0 - K_0 - f_{min}]}{T_0}, & t < t_0; \\ \frac{1}{1-t_0} \int_{t_0}^1 g(t) dt = \frac{[f_{max} + K_0 - t_0]}{1 - t_0}, & t \geq t_0. \end{cases} \quad (54)$$

Then, by the smoothing lemma shown in Appendix B, we have

$$\int_0^1 g(t) \log^2[g(t)] dt \geq \int_0^1 g_1(t) \log^2[g_1(t)] dt \equiv I \quad (55)$$

The integral evaluates to the following lower bound for the integral appearing in Eq. (53):

$$I = [t_0 - K_0 - f_{min}] \log^2 \left[\frac{t_0 - K_0 - f_{min}}{t_0} \right] + [f_{max} + K_0 - t_0] \log^2 \left[\frac{f_{max} + K_0 - t_0}{1 - t_0} \right]. \quad (56)$$

A similar argument using t_1 and K_1 gives the alternate lower bound,

$$I = [K_1 - t_1 - f_{min}] \log \left[\frac{K_1 - t_1 - f_{min}}{t_1} \right] + [f_{max} + t_1 - K_1] \log \left[\frac{f_{max} + t_1 - K_1}{1 - t_1} \right]. \quad (57)$$

In the results section we discuss the relative trade-offs between accuracy and numerical stability of these metrics as applied to backscattered waveforms acquired from molecularly targeted neovasculature.

III. MATERIALS AND METHODS

A. Nanoparticles for molecular imaging

A cross-section of the spherical liquid nanoparticles used in our study is diagrammed in Figure (6). For *in vivo* imaging we formulated nanoparticles targeted to $\alpha_v\beta_3$ -integrins of neovascularity in cancer by incorporating an “Arg-Gly-Asp” mimetic binding ligand into the lipid layer. Methods developed in our laboratories were used to prepare perfluorocarbon (perfluorooctylbromide, PFOB, which remains in a liquid state at body temperature and at the acoustic pressures used in this study²⁰) emulsions encapsulated by a lipid-surfactant monolayer.^{41,42} The Arg-Gly-Asp mimetic was coupled at a 1:1 molar ratio to malyimidophenyl-butylamide-polyethylene glycol2000-phosphatidylethanolamine resuspended from a dry lipid film in 3 mL of N_2 -purged, 6 mmol/L ethylenediaminetetraacetic acid (EDTA) by water-bath sonication for 30 minutes at 37°C to 40°C. This ligand premix was added to the remaining surfactant components, perfluorooctylbromide, and water for emulsification.⁴³ Nontargeted particles were prepared by excluding the targeting ligand. The nominal sizes for each formulation were measured with a submicron particle analyzer (Malvern Zetasizer, Malvern Instruments). Particle diameter was measured at $200 \pm 30 nm$.

B. Animal model

The study was performed according to an approved animal protocol and in compliance with guidelines of the Washington University institutional animal care and use committee.

The model used is the transgenic K14-HPV16 mouse that contains human papilloma virus-16 oncoproteins

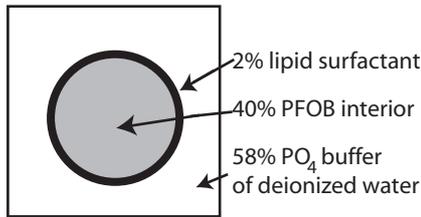


FIG. 6. A cross-sectional diagram of the nanoparticles used in our study.

driven by a keratin promoter so that lesions develop in the skin. Typically the ears exhibit squamous metaplasia, a pre-cancerous condition, associated with abundant neovasculature that expresses the $\alpha_v\beta_3$ integrin. Eight of these transgenic mice^{44,45} were treated with 1.0 mg/kg i.v. of either $\alpha_v\beta_3$ -targeted nanoparticles (n=4) or untargeted nanoparticles (n=4) and imaged dynamically for one hour using a research ultrasound imager (Vevo 660 40MHz probe) modified to store digitized RF waveforms acquired at 0, 15, 30, and 60 minute time points. All RF data were processed off-line to reconstruct images using information theoretic and conventional receivers. **Image segmentation was performed using the threshold which excluded 93% of the area under the composite histogram for all data sets (as described below in connection with Figure (9)).** The mean value of segmented pixels was computed at each time post-injection. In both control and targeted cases, the mouse was placed on a heated platform maintained at 37°C, and anesthesia was administered continuously with isoflurane gas (0.5%).

C. Ultrasonic Data Acquisition

A diagram of our apparatus is shown in Figure (7). Radio Frequency data were acquired with a research ultrasound system (Vevo 660, Visualsonics, Toronto, Canada), with an analog port and a sync port to permit digitization. The tumor was imaged with a 40 MHz single element “wobbler” probe and the RF data corresponding to single frames were stored on a hard disk for later off-line analysis. The frames (acquired at a rate of 40 Hz) consisted of 384 lines of 4096 eight-bit words acquired at a sampling rate of 500 MHz using a Gage CS82G digitizer card (connected to the analog-out and sync ports of the Vevo) in a controller PC. Each frame corresponds spatially to a region 0.8 cm wide and 0.3 cm deep.

The wobbler transducer used in this study is highly focused (3mm in diameter) with a focal length of 6 mm and a theoretical spot size of $80 \times 1100\mu\text{m}$ (lateral beam width \times depth of field at -6dB), so that the imager is most sensitive to changes occurring in the region swept out by the focal zone as the transducer is “wobbled”. Accordingly, a gel standoff was used, as shown in Figure (7), so that this region would contain the mouse ear.

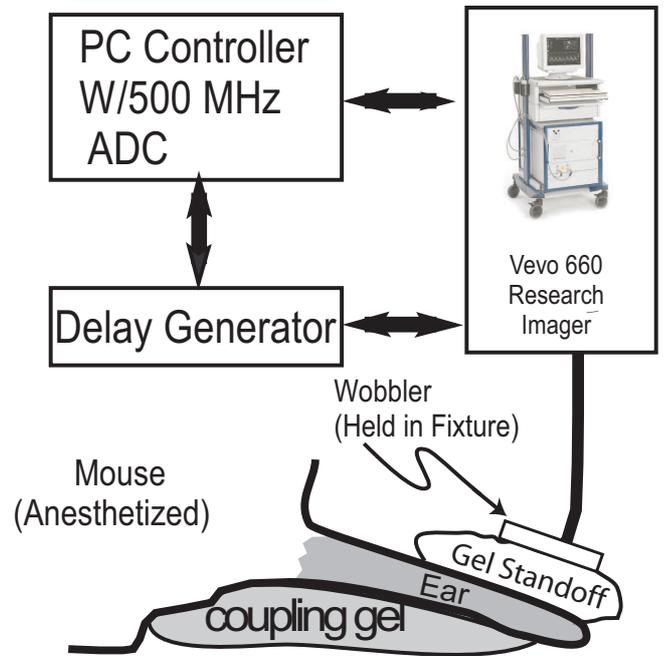


FIG. 7. A diagram of the apparatus used to acquire RF data backscattered from HPV mouse ears *in vivo*.

A close-up view showing the placement of transducer, gel standoff, and mouse ear is shown in the left side of Figure (8). Below this is a B-mode gray scale image (*i.e.* logarithm of the analytic signal magnitude). Labels indicate the location of skin (top of image insert), the structural cartilage in the middle of the ear, and a short distance below this, the echo from the skin at the bottom of the ear. To the left of this is an image of a histological specimen extracted from a HPV mouse model that has been magnified 20 times to permit better assessment of the thickness and architecture of the sites where $\alpha_v\beta_3$ targeted nanoparticle might attach (red by β_3 staining). Skin and tumor are both visible in the image. On either side of the cartilage (center band in image), extending to the dermal-epidermal junction, is the stroma. It is filled with neoangiogenic microvessels. These microvessels are also decorated with $\alpha_v\beta_3$ nanoparticles as indicated by the fluorescent image (labeled, in the upper right of the figure) of a bisected ear from an $\alpha_v\beta_3$ -injected K14-HPV16 transgenic mouse (Neumann, *et al.*, manuscript in preparation). It is in this region that the $\alpha_v\beta_3$ -targeted nanoparticles are expected to accumulate, as indicated by the presence of red β_3 stain in the magnified image of an immunohistological specimen also shown in the image.

D. Ultrasonic Data Processing

The analysis of RF frames corresponding to data in B-scan format proceeds as follows. Each of the 384 RF lines in the data was first up-sampled from 4096 to 8192

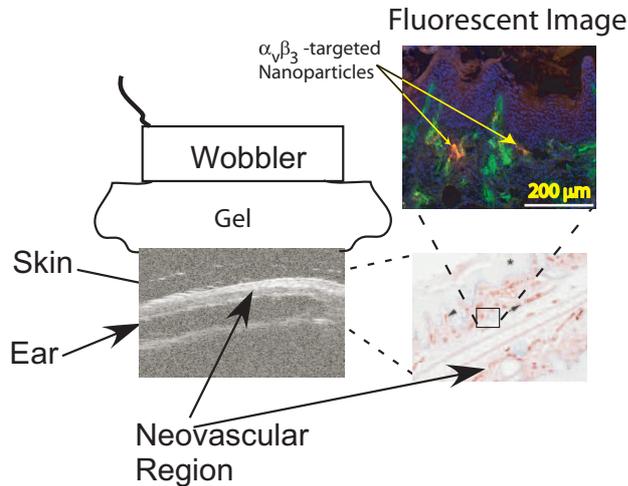


FIG. 8. Close-up of transducer, standoff and ear, with enlarged histological image showing location of binding sites, and a fluorescent image from the same anatomical region showing that $\alpha_v\beta_3$ -targeted nanoparticles indeed accumulate in this portion of the mouse ear.

points, using a cubic spline fit to the original data set in order to improve the stability of the thermodynamic receiver algorithms. Previous work has shown benefit from increased input waveform length.^{5,7} Next, a moving window analysis was performed on the upsampled data set using a rectangular window that was advanced in $0.064 \mu\text{s}$ steps (64 points), resulting in 121 window positions within the original data set. The entropy, $H_f(\log[E_f])$ within each window is used to produce the $H_f(\log[E_f])$ image.

E. Image Analysis

For this study, in which the same portion of the anatomy was imaged at successive intervals, the objective was to quantify changes in image features as a function of time. The first step in this process was the creation of a composite image from the images obtained at 0 through 60 minutes. Next, an estimate of the probability density function (PDF) of this composite image was computed by normalizing the pixel value histogram to have unit area. We emphasize that this function is not related to the density functions $w_f(y)$, that we defined in Eq. (1). Instead, it is a calculational device used to objectively segment H_f and $\log[E_f]$ images into "enhanced" and "unenhanced" regions. A typical histogram is shown in Figure (9). Its most notable feature is the presence of two local maxima; a feature that has been observed in several different experiments involving different animal models and different imaging equipment.²⁰⁻²³ The first, larger maxima, corresponds to the relatively homogeneous gray background visible in most H_f images, the smaller peak corresponds to tissue interfaces, which appear also as bright features

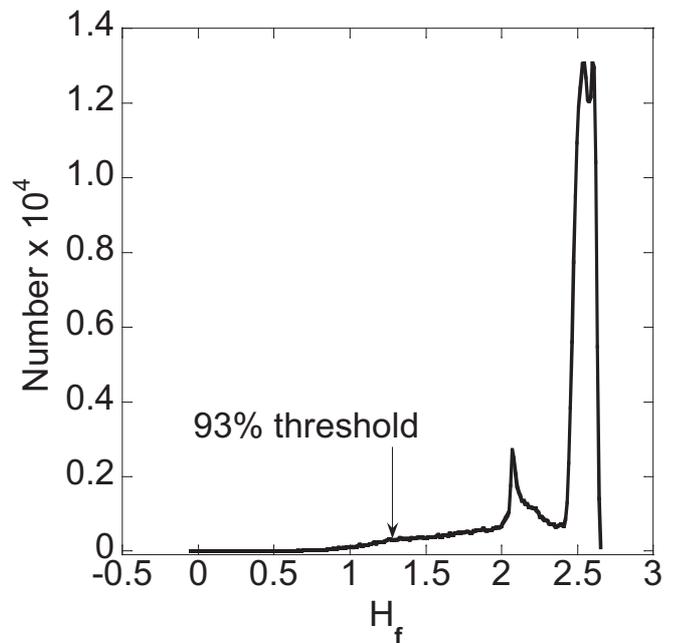


FIG. 9. A histogram from the composite H_f images acquired at 0, 15, 30, 45, and 60 minutes post-injection. The histogram has two peaks, a characteristic observed in several studies using different equipment and animal models.²¹⁻²³ The threshold level used in this study is indicated by the arrow.

in gray-scale B-mode images such as that shown in the inset of Figure (8).²²

Several different methods of image segmentation based on PDF were investigated. In all of these a specific value, or threshold, in the histogram was chosen and the images divided into two regions: those having pixel values above the threshold (considered to be unenhanced) and those having pixel values below (referred to subsequently as enhanced pixels). The PDF of all composite images exhibited a two peak structure with a large and small peak. Thresholds were set at the second minimum, and at the half way point between the large and small peaks. The full width at half maximum (FWHM) was also computed, and thresholds set at: 4.5, 3.5, 3.25, 3, 2.75, and 2.5 FWHM below the large peak. Thresholds were also set at points such that 97, 95, 93, 90, 87, and 80% of the pixel values were above the threshold. After selection of a threshold value, regions of interest (ROI) were selected using NIH ImageJ (<http://rsb.info.nih.gov/ij/>), and the mean value of the pixels lying below the threshold were computed for each of the images acquired at 0, 15, 30, 45, and 60 minutes post-injection. The mean value at zero minutes was subtracted from the values obtained for all subsequent times, to obtain a sequence of changes in receiver output as a function of time post-injection. This was done for all four animals injected with targeted nanoparticles and for the four control animals. These sequences of relative changes were then averaged over targeted and control groups to obtain a sequence of time

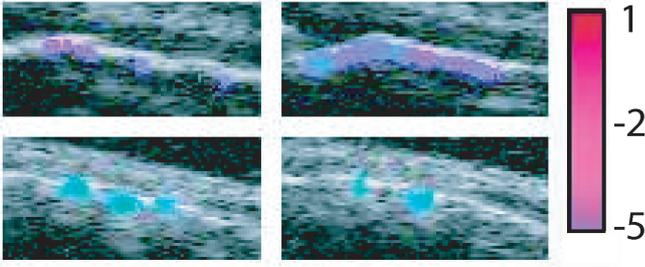


FIG. 10. Cropped images showing a comparison of H_f enhanced conventional images showing the effects of targeting (top row) vs. control (bottom row).

points for change in receiver output for both groups of animals. The threshold of 93% was finally chosen since it produced the smallest p-value (0.00043) for a t-test comparing the mean values of the ROI at 15 minutes as compared to 60 minutes. The corresponding p-value for the control group was 0.27.

IV. RESULTS

In order to present the data in a format that combines the superior spatial resolution of conventional medical imaging with the superior detection capability of the entropy receiver we have employed a scheme similar to Doppler-enhanced imaging. Namely, we color-code pixels of the conventional grey-scale composite image using the corresponding pixel in the entropy image, to produce a composite image like that shown in Figure (10). This figure was made by coloring the entropy image and then superposing the colored pixels lying below the 93% threshold onto the conventional image to indicate regions of “enhancement”.

The top panel of Figure (11) compares the growth, with time post-injection, of the mean value of the enhanced regions of H_f images obtained from all eight of the animals used in our study. Standard error bars are shown with each point. These data were obtained by computing the mean value of pixels lying below the 93% threshold at each time point (0, 15, 30, 45, and 60 minutes) for each animal (4 injected with targeted nanoparticles, 4 injected with nontargeted nanoparticles) as discussed above. As the data show, the mean value, or enhancements, obtained in the targeted group increase linearly with time. After 30 minutes the mean value of enhancement is measurable. Moreover, the values at 15 and 60 minutes are statistically different ($p < 0.00043$). The bottom panel of the figure shows the corresponding result obtained from the control group of animals that were injected with nontargeted nanoparticles. There is no discernible trend in the group and the last three time points are not statistically different from zero. Comparison of the enhancement measured at 15 and 60 minutes

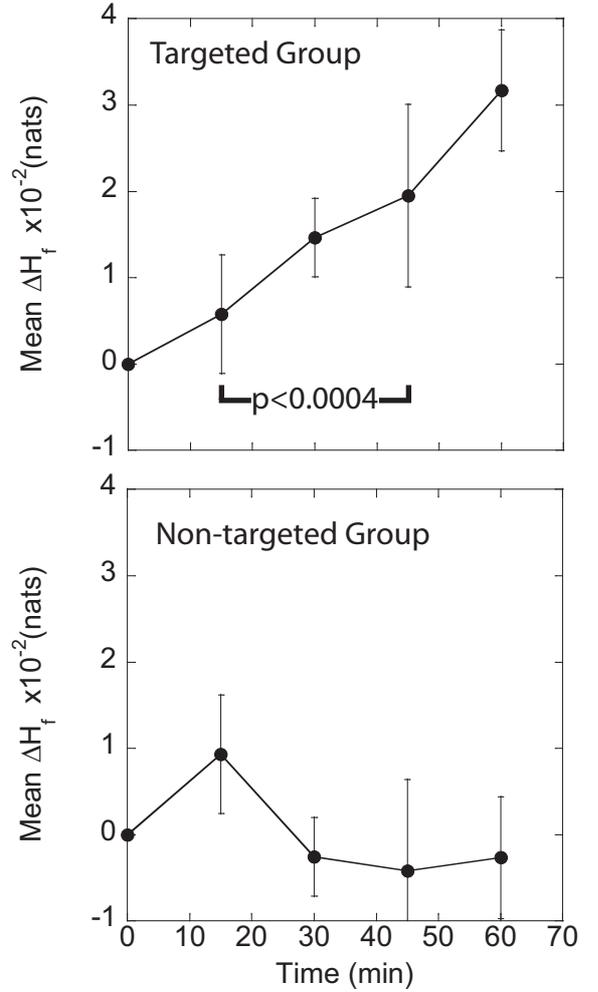


FIG. 11. A plot of average enhancement obtained by analysis of H_f images from nine HPV mice injected with $\alpha_v\beta_3$ targeted nanoparticles and five control HPV mice given no injection.

yields a p-value ≈ 0.27 .

The corresponding results obtained using the $\log[E_f]$ receiver are shown in Figure (12), whose top panel compares the growth, with time, of the mean value of the enhanced regions of $\log[E_f]$ images obtained from all eight of the animals used in our study. Standard error bars are shown with each point. These data were obtained by computing the mean value of pixels lying below the 93% threshold at each time point (0, 15, 30, 45, and 60 minutes) for each animal (4 injected with targeted nanoparticles, 4 injected with nontargeted nanoparticles) as discussed above. Unlike the entropy case, the values at 15 and 60 minutes are not statistically different ($p = 0.10$). The bottom panel of the figure shows the corresponding result obtained from the control group of animals that were injected with nontargeted nanoparticles. There is no discernible trend in the group and the last three time points are not statistically different from zero.

The results summarized in Figures (11) and (12) are

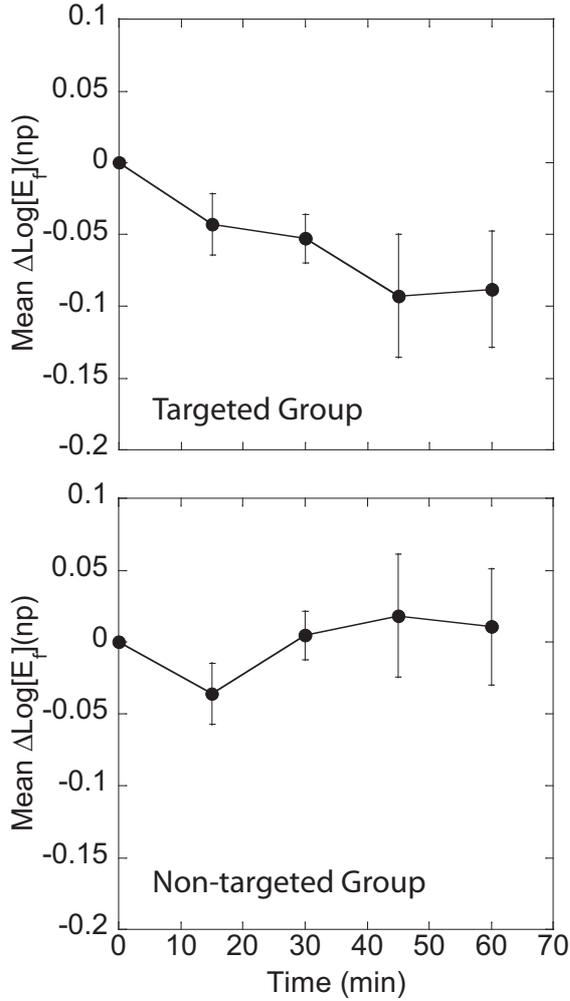


FIG. 12. A plot of average enhancement obtained by analysis of $\log[E_f]$ images from nine HPV mice injected with $\alpha_v\beta_3$ targeted nanoparticles and five control HPV mice given no injection.

consistent with Eqs. (47), (50), (51), (56) and (57). This is demonstrated in Figure (13), which uses these equations to estimate the relative sensitivities, defined as the percentage change of receiver output. To obtain this fraction, the following images were made using the same 15 minute data sets that were used to make Figure (11): maximum average relative receiver sensitivity magnitude for $\log E_f$ (Eq. (47)), minimum average relative receiver sensitivity magnitude, $\|\log[w_f(y)]^2\|_2$, for H_f (Eq. (45)) with $\|y^2\|_2$, the “minimum $w_f(y)$ estimate” (Eq. (50)), the “maximum $f'_\uparrow(t)$ estimate” (Eq. (51)), K_0 (Eq. (56)), and K_1 (Eq. (57)), respectively. To these were applied the ROI defined by the H_f images thresholded at the 93% level. Subsequently, the mean pixel value within the ROI was computed. Having the numerator of the sensitivity, the same ROI was applied to the $\log[E_f]$ image, and the H_f image respectively, to compute a reference value (*i.e.*, denominator) for either the sensitivity of $\log[E_f]$ or H_f

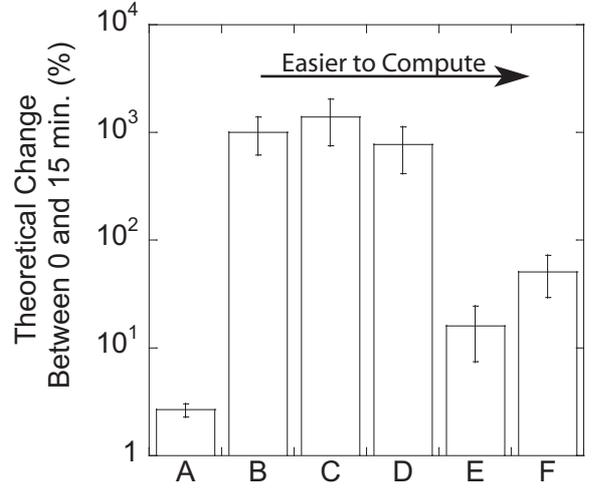


FIG. 13. Comparisons of receiver sensitivities, as defined in the text, for 15 minute data sets shown in Figure (10).

images. These values are plotted as a percentage change relative to the 15 minute value in Figure (13), so that lower values indicate lower relative sensitivity. As the figure shows (Column A) $\log[E_f]$ has the lowest sensitivity, while the sensitivity for H_f , column B, is high, consistent with the results of Figure (11). Column C shows the estimate of $|\Delta H_f|$ based on the “minimum $w_f(y)$ estimate” (Eq. (50)). The mean value provided by this estimate is apparently a little greater than the “exact” value although the error bar for this estimate is also greater. These result from use of a single point from $w_f(y)$, which is susceptible to Fourier “ringing”, to obtain column C versus the integration of a function of $w_f(y)$, to obtain column B. However, the error bars of columns B and C overlap, so the estimates agree to within experimental error. Column D shows the estimate of entropy sensitivity obtained using the “maximum $f'_\uparrow(t)$ estimate” (Eq. (51)). This estimate, which has the advantage of being rapidly computable in the time domain but the disadvantage of requiring the potentially noise enhancing operation of numerically differentiating an experimental waveform, also agrees, to within experimental error, with columns B and C. Estimates that do not require numerical differentiation of experimental data are shown in columns E and F. These were obtained using K_0 and K_1 (Eqs. (56) and (57)) to compute the numerator of the sensitivity fraction and are not susceptible to either Fourier “ringing” or noise enhancement by numerical differentiation. They are thus more stable than the estimates in columns C and D. All of these estimates successfully predict the greater relative sensitivity of H_f versus $\log[E_f]$ imaging shown in Figure (11). The label on the plot indicates that they appear, from left to right, approximately in order of decreasing computational cost and increasing numerical stability, with the two columns, E, F, being computed in the time domain after the relatively inexpensive step

of sorting the data. As these bars indicate, it is possible to assess, *a priori*, the sensitivity of entropy imaging without having to compute it. This information could be very useful in imaging applications, which must balance computing requirements and frame-rate constraints

V. DISCUSSION

The fact that the results displayed in Figure (13) could have been used to predict the outcome of our experimental study (as summarized in Figures (11) and (12)) indicates the utility of the qualitative benchmarks derived in Section II (Approach). This suggests the existence of a collection of performance benchmarks, useful as an intuitive guide for experimental application of entropy receivers, similar to that which already exists for more conventional signal analysis. Traditional B-mode grayscale imaging (which is $\log[E_f]$ imaging in the limit where the “moving window” or “box-car” is the same size as the digitization interval⁴⁶) is a specific example where intuitive ideas about the effects of changing gain settings, attenuation from intervening tissue, noise, or introduction of bubble-based contrast agents are routinely applied in both the laboratory and clinic.

We have described the impact of some of these factors on the entropy-based receiver, H_f . In sections II.E, we showed that H_f is invariant with respect to rescaling of $f(t)$. In section II.G.3 and II.G.5, we derived an average-case relationship between the sensitivities of H_f and energy-based receivers. The utility of this relation is demonstrated by our experimental results.

We have provided, in Section IIC, a preliminary, although indirect, step towards a physical interpretation for H_f in the case of scattering problems. In that section, we described a simplified scattering model and indicated its relation to $w_f(y)$. We observed that different scattering model configurations could lead to the same density for the backscattered waves, which implies they would have the same entropies and same energies. However, from an experimental point of view the most important characteristic of any receiver is its sensitivity to subtle changes in scatterer configuration. For the study described here the entropy is more sensitive.

Ultimately, we desire a physical interpretation of H_f similar to that for energy. In acoustics, perhaps the simplest interpretation of energy arises from spring models employed to describe the interaction of sound with the supporting medium. Deeper insight is obtained by application of Hamiltonian or Lagrangian mechanics to modeling of wave propagation and scattering in a supporting medium. This setting provides immediate access to a thoroughly developed concept of energy that has proven predictive power. At this time there is no corresponding body of knowledge for H_f , and although initial experiments^{3-7,20-24} have been promising and we have taken initial steps toward understanding its physical basis, the ultimate scope of applicability for this quantity

remains to be determined.

APPENDIX A: TWO IMPLICATIONS OF STATISTICAL INDEPENDENCE

Even though the concept of independence is widely used, it can have surprising implications when considered from the frame of mind usually adopted in deterministic signal processing. We provide two examples below. In the first we show that there do not exist a complete orthonormal set of functions that are statistically independent. In the second we show that there is no measurable function (*i.e.*, random variable) that is statistically independent of the waveform $f(t) = t$.

1. Statistical independence vs. linear independence of functions

In this appendix we discuss statistical independence and its relation to eigenfunction-based signal processing. The definition of a density function $w_f(y)$ corresponds to P_X where X is a random variable. In the physics literature the term random variable is associated with functions that are unpredictable, and thus infinitely rough (*i.e.*, everywhere continuous but nowhere differentiable), such as the Brownian motion. However, the mathematical definition of a “random” variable consists solely of the requirement that a function be measurable.^{47,48} Hence, familiar functions such as $\sin(t)$ and e^t are also random variables (once a domain, *e.g.*, $[0, 1]$ is chosen), with well-defined density functions. In probabilistic notation the density function $w_f(y)$ of $f(t)$ would be written as P_f . Just as two random functions X and Y have a joint density function $P_{X,Y}$ so do functions like $\sin(t)$ and e^t . In our notation, which is borrowed from measure theory,¹¹ the joint density function of two waveforms, $y = f(t)$ and $x = g(t)$, would be written $w_{f,g}(y, x)$. The fundamental integral relationship, Eq. (1), becomes

$$\int_0^1 \phi(f(t), g(t)) dt = \int_0^1 \int_0^1 \phi(y, x) w_{f,g}(y, x) dy dx. \quad (\text{A1})$$

The notion that two functions $f(t)$, $g(t)$ are independent is expressed in the measure-theoretic notation by the equation $w_f(y)w_g(x) = w_{f,g}(y, x)$. To see that this is a very strong property, possessed by few of the functions of ordinary calculus, let $\phi(y, x) = yx$ in Eq. (A1),

then

$$\begin{aligned}\int_0^1 \phi(f(t), g(t)) dt &= \int_0^1 \int_0^1 \phi(y, x) w_{f,g}(y, x) dy dx \\ &= \int_0^1 \int_0^1 y x w_f(y) w_g(x) dy dx \\ &= \int_0^1 y w_f(y) dy \int_0^1 x w_g(x) dx \\ &= \int_0^1 f(t) dt \int_0^1 g(t) dt,\end{aligned}$$

that is,

$$\int_0^1 f(t) g(t) dt = \int_0^1 f(t) dt \int_0^1 g(t) dt, \quad (\text{A2})$$

which is certainly not true in general. To see how remarkable the notion of independence really is, assume that $f(t)$ has zero mean, then for all $g(t)$ independent of $f(t)$

$$\int_0^1 f(t) g(t) dt = 0 \quad (\text{A3})$$

Moreover, since Eq. (A2) generalizes immediately to

$$\int_0^1 f(t) \phi(g(t)) dt = \int_0^1 f(t) dt \int_0^1 \phi(g(t)) dt, \quad (\text{A4})$$

we see that $f(t)$ and any continuous function, ϕ , of $g(t)$ are orthogonal.

Among other things, this implies that there are no complete orthonormal sets of statistically independent functions.

To see this, suppose, on the contrary, that there exist a complete set of orthonormal functions, $\{\psi_k(t) | k = 1, \dots\}$, which are pairwise statistically independent. For mathematical simplicity we will also assume that these functions are real, bounded and that any three of the $\psi_k(t)$ are independent.

Since the set is complete we must be able to represent a constant using this basis so that for some k ,

$$\int_0^1 \psi_k(t) dt \neq 0. \quad (\text{A5})$$

Without loss of generality, we may reorder the indexing of the $\psi_k(t)$ so that this happens for $k = 1$.

If $j \neq 1$, then by orthogonality and independence,

$$0 = \int_0^1 \psi_1(t) \psi_j(t) dt = \int_0^1 \psi_1(t) dt \int_0^1 \psi_j(t) dt, \quad (\text{A6})$$

So that

$$\int_0^1 \psi_j(t) dt = 0, \quad (\text{A7})$$

for all $k > 1$.

Since the $\psi_k(t)$ are bounded, any power of $\psi_k(t)$ is integrable. Thus, we observe that for every $k \neq j, 1$

$$\int_0^1 \psi_j(t)^2 \psi_k(t) dt = \int_0^1 \psi_j(t)^2 dt \int_0^1 \psi_k(t) dt = 0, \quad (\text{A8})$$

For $j = 1$, this implies that all coefficients except the first are zero, and so $\psi_1^2(t) = \alpha \psi_1(t)$ for some constant α , *i.e.*, $\psi_1(t) = \alpha$. Moreover, $|\alpha| = 1$, since the $\psi_k(t)$ have norm 1.

For $j > 1$, Eq. (A8) tells us that all coefficients except the 1st and j^{th} are zero. Thus, the expansion for $\psi_j^2(t)$ is

$$\psi_j^2(t) = a_j + b_j \psi_j(t), \quad (\text{A9})$$

i.e., every value of $\psi_j(t)$ satisfies the same quadratic equation. Hence, $\psi_j(t)$ can have only two distinct values, which we will denote by a_j and a'_j .

Letting $\mathbb{I}_{A_j}(t)$ and $\mathbb{I}_{A'_j}$ denote the characteristic functions of the sets A_j and A'_j , where $\psi_j(t)$ takes on the values a_j and a'_j respectively, we write

$$\psi_j(t) = a_j \mathbb{I}_{A_j}(t) + a'_j \mathbb{I}_{A'_j}(t). \quad (\text{A10})$$

Specifically, for $j = 2$, we have

$$\psi_2(t) = a_2 \mathbb{I}_{A_2}(t) + a'_2 \mathbb{I}_{A'_2}(t). \quad (\text{A11})$$

We will now construct a non-zero function that cannot be represented in terms of the $\psi_j(t)$, contradicting the assumption of their completeness. Consider

$$f(t) = \psi_3(t) \mathbb{I}_{A_2}(t). \quad (\text{A12})$$

We observe that any bounded measurable function $\phi(t)$ satisfying $\phi(a_2) = 1$ and $\phi(a'_2) = 0$ also satisfies

$$\mathbb{I}_{A_j}(t) = \phi(\psi_2(t)). \quad (\text{A13})$$

So, by pairwise independence,

$$\begin{aligned}\int_0^1 f(t) \psi_1(t) dt &= \int_0^1 \psi_3(t) \mathbb{I}_{A_2}(t) dt \\ &= \int_0^1 \psi_3(t) \phi(\psi_2(t)) dt \\ &= \int_0^1 \psi_3(t) dt \int_0^1 \phi(\psi_2(t)) dt \\ &= 0,\end{aligned} \quad (\text{A14})$$

and

$$\begin{aligned}\int_0^1 f(t) \psi_2(t) dt &= \int_0^1 \psi_3(t) \mathbb{I}_{A_2}(t) \psi_2(t) dt \\ &= \int_0^1 \psi_3(t) \phi(\psi_2(t)) \psi_2(t) dt \\ &= \int_0^1 \psi_3(t) dt \int_0^1 \phi(\psi_2(t)) \psi_2(t) dt \\ &= 0,\end{aligned} \quad (\text{A15})$$

since $\int_0^1 \psi_3(t)dt = 0$. For $k > 3$, we have, using, triple-wise independence,

$$\begin{aligned} \int_0^1 f(t)\psi_2(t)dt &= \int_0^1 \psi_3(t)\mathbb{I}_{A_2}(t)\psi_k(t)dt & (A16) \\ &= \int_0^1 \psi_3(t)\phi(\psi_2(t))dt\psi_k(t) \\ &= \int_0^1 \psi_3(t)dt \int_0^1 \phi(\psi_2(t)) \int_0^1 \psi_k(t)dt \\ &= 0. \end{aligned}$$

Thus

$$f(t) = \alpha\psi_3(t), \quad (A17)$$

or

$$\psi_3(t)\mathbb{I}_{A_2}(t) = \alpha\psi_3(t), \quad (A18)$$

This implies that $\psi_3(t) = 0$ for any t in $[0, 1]$ that is not in A_2 , *i.e.*, it takes on 0 as one of its (two possible) values contradicting $\int_0^1 \psi_3(t)dt = 0$. Thus, it is not possible to expand $f(t)$ in terms of the $\psi_k(t)$, which contradicts the assumption that they comprise a complete set of functions. Specific examples of independent functions may be found in the articles by Agnew and Kac, Karlin, and Kuipers.¹⁷⁻¹⁹

2. Statistical independence and the function $f(t) = t$

In this section we prove the somewhat surprising result that there does not exist a non-constant function that is statistically independent of the function $y = f(t) = t$. This result may be generalized to other monotonic functions, as discussed at the end of this section. The linear case is presented since it is the simplest case and illustrates the basic ideas in the most straightforward way.

We begin with the definition of the essential range, R_f , of a function $f(t)$, which is the set of all real numbers, z such that $\{x : |f(x) - z| < \epsilon\}$ has positive measure for all $\epsilon > 0$. Expressed in terms of probability, this can be thought of as the set of all points, z , such that for any neighborhood of z , $f(t)$ takes on values in that neighborhood with non zero probability, or if $z \in I$, where I is an open interval then, $f^{-1}(I)$ has positive measure, *i.e.* $|f^{-1}(I)| > 0$ where we have used the symbol $|S|$ to denote the measure of a set S as we will do for the remainder of this section. In terms of digital waveforms, which may have sharp transients, R_f can be thought of as the range of $f(t)$ without the transients as shown in Figure (14).

A function whose essential range is only one point is said to be essentially constant. We claim that there does not exist a measurable function $g(t)$ defined on $[0, 1]$ that is statistically independent of $f(t)$, unless $g(t)$ is essentially constant.

To prove this assume that $g(t)$ is not essentially constant, *i.e.*, that the essential range of $g(t)$ has at least

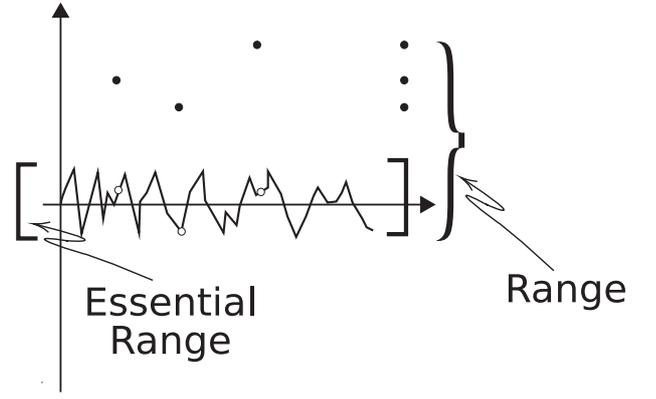


FIG. 14. A diagram of a noisy time-domain signal. The full range of the signal includes the transients. However, the essential range, indicated by the bracket beside the vertical axis excludes these.

two points. Let $\alpha, \beta, \alpha \neq \beta$, be in the essential range of $g(t)$. Let $\alpha \in I_\alpha, \beta \in I_\beta, I_\alpha, I_\beta$ open intervals with $I_\alpha \cap I_\beta = \emptyset$ (pg. 98 of Munkres⁴⁹).

Now, since $f(t)$ and $g(t)$ are independent, for any measurable sets, A, B ,

$$|f^{-1}(A) \cap g^{-1}(B)| = |f^{-1}(A)| \times |g^{-1}(B)| \quad (A19)$$

i.e. using probabilistic notation

$$\mathbb{P}[f \in A \text{ and } g \in B] = \mathbb{P}[f \in A] \times \mathbb{P}[g \in B] \quad (A20)$$

Let $A = g^{-1}(I_\alpha), B = I_\beta$. Then the assumption of independence leads to

$$|f^{-1}(g^{-1}(I_\alpha)) \cap g^{-1}(I_\beta)| = |f^{-1}(g^{-1}(I_\alpha))| \times |g^{-1}(I_\beta)| \quad (A21)$$

Now $f^{-1}(g^{-1}(I_\alpha)) = g^{-1}(I_\alpha)$, since $f(t) = t$, so

$$\begin{aligned} |f^{-1}(g^{-1}(I_\alpha)) \cap g^{-1}(I_\beta)| &= |g^{-1}(I_\alpha) \cap g^{-1}(I_\beta)| & (A22) \\ &= |g^{-1}(I_\alpha \cap I_\beta)| = 0 \end{aligned}$$

On the other hand

$$\begin{aligned} |f^{-1}(g^{-1}(I_\alpha)) \cap g^{-1}(I_\beta)| & & (A23) \\ &= |f^{-1}(g^{-1}(I_\alpha))| \times |g^{-1}(I_\beta)| \\ &= |g^{-1}(I_\alpha)| \times |g^{-1}(I_\beta)| \\ &\neq 0 \end{aligned}$$

since $|g^{-1}(I_\alpha)| \neq 0$ and $|g^{-1}(I_\beta)| \neq 0$ by the definition of the essential range. If we let $A = f(g^{-1}(I_\alpha)), B = I_\beta$, where $f(t)$ is any monotonic function, then the discussion above is essentially unchanged and we see that result may be generalized to the case of monotonic functions.

APPENDIX B: SMOOTHING DECREASES ENTROPY

The following lemma states that smoothing a density function decreases its entropy. Let

$$g(t) = \frac{d}{dt}[f_{\uparrow}(t)]. \quad (\text{B1})$$

If $g > 0$ on any interval $I \subset [0, 1]$ and we define

$$g_1(t) = \begin{cases} g(t), & t \notin I; \\ \frac{1}{|I|} \int_0^1 g(t) dt & t \in I, \end{cases} \quad (\text{B2})$$

where $|I|$ denotes the measure of the set I , then

$$\int_0^1 g_1(t) \log_2[g_1(t)] dt < \int_0^1 g(t) \log_2[g(t)] dt \quad (\text{B3})$$

$$\int_0^1 g_2(t) \log_2[g_2(t)] dt - \int_0^1 g(t) \log_2[g(t)] dt = \varepsilon \int_{E_1} (\log_2[g_2(t)] + 2 \log[g(t)]) dt - \varepsilon \int_{E_2} (\log_2[g_2(t)] + 2 \log[g(t)]) dt + O[\varepsilon^2] \quad (\text{B6})$$

As $g(t)$ on E_1 is everywhere less than $g(t)$ on E_2 , this is negative to first order. So we can reduce the entropy integral by smoothing $g(t)$.

ACKNOWLEDGMENTS

This study was funded by NIH EB002168, HL042950, and CO-27031 and NSF DMS 0501079. The research was carried out at the Washington University Department of Mathematics and the School of Medicine.

- ¹ A. H. Schmieder, P. M. Winter, S. D. Caruthers, T. D. Harris, T. A. Williams, J. S. Allen, E. K. Lacy, H. Zhang, M. J. Scott, G. Hu, J. D. Robertson, S. A. Wickline, and G. M. Lanza, "Molecular mr imaging of melanoma angiogenesis with $\alpha_v\beta_3$ -targeted paramagnetic nanoparticles", *Magnetic Resonance in Medicine* **53**, 621–627 (2005).
- ² C. S. Hall, J. N. Marsh, M. J. Scott, P. J. Gaffney, S. A. Wickline, and G. M. Lanza, "Time evolution of enhanced ultrasonic reflection using a fibrin-targeted nanoparticulate contrast agent", *Journal of the Acoustical Society of America* **108**, 3049–3057 (2000).
- ³ M. Hughes, "A comparison of shannon entropy versus signal energy for acoustic detection of artificially induced defects in plexiglass", *Journal of the Acoustical Society of America* **91**, 2272–2275 (1992).
- ⁴ M. Hughes, "Analysis of ultrasonic waveforms using shannon entropy", *Proceedings of the 1992 I.E.E.E. Ultrasonics Symposium* **2**, 1205–1209 (1992).
- ⁵ M. Hughes, "Analysis of digitized waveforms using shannon entropy", *Journal of the Acoustical Society of America* **93**, 892–906 (1993).

unless $g_1 \equiv g(t)$.

Proof : If g is not constant on I , choose two sets $E_1, E_2 \subset I$ such that $|E_1| = |E_2| > 0$, and

$$g|_{E_1} < \frac{1}{|I|} \int_0^1 g(t) dt - 2\varepsilon \quad (\text{B4})$$

$$g|_{E_2} > \frac{1}{|I|} \int_0^1 g(t) dt + 2\varepsilon \quad (\text{B5})$$

for some $\varepsilon > 0$. Let $g_2 = g + \varepsilon\chi_{E_1} - \varepsilon\chi_{E_2}$, where χ_{E_i} is the characteristic function of the set E_i . Then

- ⁶ M. Hughes, "Nde imaging of flaws using rapid computation of shannon entropy", *Proceedings of the 1993 I.E.E.E. Ultrasonics Symposium* **2**, 697–700 (1993).
- ⁷ M. Hughes, "Analysis of digitized waveforms using shannon entropy. ii. high-speed algorithms based on green's functions", *Journal of the Acoustical Society of America* **95**, 2582–2588 (1994).
- ⁸ R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance* (Chelsea Publishing Company, New York, NY) (1987).
- ⁹ N. Wiener, *Extrapolation, interpolation, and smoothing of stationary time series : with engineering applications* (M.I.T. Press, Cambridge, MA) (1949).
- ¹⁰ U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series* (Chelsea Publishing Company, New York, NY) (1984).
- ¹¹ R. L. Wheeden and A. Zygmund, *Measure and Integral: An Introduction to Real Analysis* (Marcel-Dekker, New York) (1977).
- ¹² T. M. Cover and J. A. Thomas, *Elements of Information Theory* (wiley-Interscience, New York) (1991).
- ¹³ S. Kullback, *Information theory and statistics* (Dover, New York) (1997).
- ¹⁴ F. M. Reza, *An Introduction to Information Theory* (Dover, New York) (1994).
- ¹⁵ L. J. Ziomek, *Fundamentals of Acoustic Field Theory and Space-Time Signal Processing* (CRC Press, Boca Raton) (1995).
- ¹⁶ A. Soklakov, "Occam's razor as a formal basis for a physical theory", *Foundations of Physics Letters* **15**, 107–35 (2002).
- ¹⁷ K. Agnew and M. Kac, "Translated functions and statistical independence", *Bulletin of the American Mathematical Society* **7**, 148–154 (1941).
- ¹⁸ S. Karlin, "Orthogonal properties of independent func-

- tions”, Transactions of the American Mathematical Society **66**, 44–64 (1949).
- 19 L. Kuipers, “Continuous distribution mod 1 and independence of functions”, Nieuw Archief voor Wiskunde **XI**, 1–3 (1963).
 - 20 M. S. Hughes, J. N. Marsh, J. Arbeit, R. Neumann, R. W. Fuhrhop, G. M. Lanza, and S. A. Wickline, “Ultrasonic molecular imaging of primordial angiogenic vessels in rabbit and mouse models with $\alpha_v\beta_3$ -integrin targeted nanoparticles using information-theoretic signal detection: Results at high frequency and in the clinical diagnostic frequency range”, Proceedings of the 2005 I.E.E.E. Ultrasonics Symposium (2005).
 - 21 M. S. Hughes, J. N. Marsh, C. S. Hall, D. Savy, M. J. Scott, J. S. Allen, E. K. Lacy, C. Carradine, G. M. Lanza, and S. A. Wickline, “In vivo ultrasonic detection of angiogenesis with site-targeted nanoparticle contrast agents using measure-theoretic signal receivers”, Proceedings of the 2004 I.E.E.E. Ultrasonics Symposium **2**, 1106 – 1109 (2004).
 - 22 M. S. Hughes, J. N. Marsh, C. S. Hall, D. Savy, M. J. Scott, J. S. Allen, E. K. Lacy, C. Carradine, G. M. Lanza, and S. A. Wickline, “Characterization of digital waveforms using thermodynamic analogs: Applications to detection of materials defects”, I.E.E.E Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **52**, 1555 – 1564 (2005).
 - 23 M. Hughes, J. Marsh, A. Woodson, E. Lacey, C. Carradine, G. M. Lanza, and S. A. Wickline, “Characterization of digital waveforms using thermodynamic analogs: Detection of contrast targeted tissue in mda 435 tumors implanted in athymic nude mice”, Proceedings of the 2005 I.E.E.E. Ultrasonics Symposium (2005).
 - 24 M. S. Hughes, J. N. Marsh, H. Zhang, A. K. Woodson, J. S. Allen, E. K. Lacy, C. Carradine, G. M. Lanza, and S. A. Wickline, “Characterization of digital waveforms using thermodynamic analogs: Detection of contrast-targeted tissue in vivo”, I.E.E.E Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **53**, 1609–1616 (2006).
 - 25 A. S. Weigend and N. A. Gershenfeld, *Time Series Prediction* (Addison-Wesley, Reading, MA) (1994).
 - 26 E. E. Merzbacher, J. Feagin, and T. Wu, “Superposition of the radiation from n independent sources and the problem of random flights”, American Journal of Physics **45**, 964 (1977).
 - 27 C. Burckhardt, “Speckle in ultrasound b-mode scans”, .
 - 28 R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, “Statistics of speckle in ultrasound b-scans”, I.E.E.E. Transactions on Ultrasonics Ferroelectrics and Frequency Control **30**, 156–163 (1983).
 - 29 T. A. Tuthill, R. H. Sperry, and P. K. J, “Deviations from rayleigh statistics in ultrasonic speckle”, Ultrasonic Imaging **10**, 81–9 (1988).
 - 30 L. Weng, J. M. Reid, P. M. Shankar, and K. Soetanto, “Ultrasound speckle analysis based on the k distribution,” .
 - 31 R. C. Molthen, P. M. Shankar, and J. M. Reid, “Characterization of ultrasonic b-scans using non-rayleigh statistics”, Ultrasound in Medicine and Biology **21**, 161–70 (1995).
 - 32 V. Dutt and J. F. Greenleaf, “Statistics of the log-compressed echo envelope”, Journal of the Acoustical Society of America **99**, 3817–25 (1996).
 - 33 P. M. Shankar, R. Molthen, V. M. Narayanan, J. M. Reid, V. Genis, F. Forsberg, C. W. Piccoli, A. E. Lindenmayer, and B. B. Goldberg, “Studies on the use of non-rayleigh statistics for ultrasonic tissue characterization”, Ultrasound in Medicine and Biology **22**, 873–82 (1996).
 - 34 P. M. Shankar, “A compound scattering pdf for the ultrasonic echo envelope and its relationship to k and nakagami distributions”, I.E.E.E. Transactions on Ultrasonics Ferroelectrics and Frequency Control **50**, 339–43 (2003).
 - 35 P. Tsui and S. Wang, “The effect of transducer characteristics on the estimation of nakagami parameter as a function of scatterer concentration”, Ultrasound in Medicine and Biology **30**, 1345–53 (2004).
 - 36 E. Ashton and K. Parker, “Multiple resolution bayesian segmentation of ultrasound images,” .
 - 37 G. Cloutier, M. Daronatand, D. Savery, D. Garcia, L. G. Durand, and F. Foster, “Non-gaussian statistics and temporal variations of the ultrasound signal backscattered by blood at frequencies between 10 and 58 mhz”, Journal of the Acoustical Society of America **116**, 566–77 (2004).
 - 38 R. C. Molthen, P. M. Shankar, J. M. Reid, F. Forsberg, E. J. Halpern, C. W. Piccoli, and B. B. Goldberg, “Comparisons of the rayleigh and k-distribution models using in vivo breast and liver tissue”, Ultrasound in Medicine and Biology **24**, 93–100 (1998).
 - 39 P. Shankar, “A general statistical model for ultrasonic backscattering from tissues”, I.E.E.E. Transactions on Ultrasonics Ferroelectrics and Frequency Control **47**, 727–736 (2000).
 - 40 P. Beckman and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces* (Macmillan Co.).
 - 41 S. Flacke, S. Fischer, M. J. Scott, R. J. Fuhrhop, J. S. Allen, M. McLean, P. Winter, G. A. Sicard, P. J. Gaffney, S. A. Wickline, and G. M. Lanza, “Novel mri contrast agent for molecular imaging of fibrin implications for detecting vulnerable plaques”, Circulation **104**, 1280–1285 (2001).
 - 42 G. M. Lanza, K. D. Wallace, M. J. Scott, W. P. Cacheris, D. R. Abendschein, D. H. Christy, A. M. Sharkey, J. G. Miller, P. J. Gaffney, and S. A. Wickline, “A novel site-targeted ultrasonic contrast agent with broad biomedical application”, Circulation **94**, 3334–3340 (1996).
 - 43 P. M. Winter, S. D. Caruthers, X. Yu, S. K. Song, J. J. Chen, B. Miller, J. W. M. Bulte, J. D. Robertson, P. J. Gaffney, S. A. Wickline, and G. M. Lanza, “Improved molecular imaging contrast agent for detection of human thrombus”, Magnetic Resonance in Medicine **50**, 411–416 (2003).
 - 44 J. M. Arbeit, R. R. Riley, B. Huey, C. Porter, G. Kelloff, R. Lubet, J. M. Ward, and D. Pinkel, “chemoprevention of epidermal carcinogenesis in k14-hpv16 transgenic mice”, Cancer Research **59**, 3610–3620 (1999).
 - 45 J. M. Arbeit, K. Mnger, P. M. Howley, and D. Hanahan, “Progressive squamous epithelial neoplasia in k14-human papillomavirus type 16 transgenic mice”, Journal of Virology **68**, 4358–4368 (1994).
 - 46 L. Ziomek, *Fundamentals of Acoustic Field Theory and Space-Time Signal Processing* (CRC Press, Boca Raton) (1995).
 - 47 A. Kolmogorov, *Foundations of the Theory of Probability* (Chelsea Publishing Company, New York) (1950).
 - 48 G. B. Folland, *Real Analysis Modern Techniques and Their Applications* (John Wiley and Sons, Inc.).
 - 49 J. R. Munkres, *Topology* (Prentice Hall, Upper Saddle River, NJ) (2000).