# Thz Medical Imaging

**Akash K Singh, PhD**
IBM Corporation Sacramento, USA

**Abstract**

**The terahertz (THz) part of the electromagnetic spectrum falls between the lower frequency millimetre wave region and, at higher frequencies, the far-infrared region. The frequency range extends from 0.1 THz to 10 THz, where both these limits are rather loose. As the THz region separates the more established domains of microwaves and optics, a typical THz technique will incorporate aspects of both realms, and may even draw on the best of both. The two bounding parts of the spectrum also yield distinct sets of methods of generating and detecting THz waves. These approaches can thus be categorised as having either microwave or optical/photonic origins. As a result of breakthroughs in technology, the THz region is finally finding applications outside its traditional heartlands of remote sensing and radio astronomy. Extensive research has identified many attractive uses and has paved the technological path towards flexible and accessible THz systems. Examples of novel applications include medical and dental imaging, gene theory, communications and detecting the DNA sequence of virus and bacteria. The presentation will discuss the range of THz applications and will present the components and systems that are utilised for the frequency region.**

**Keywords- Terahertz (THz), Medical Imaging, and Infrared**

## I. INTRODUCTION

Although the terahertz (THz) field first used the moniker "THz" to describe research in the 100 GHz to 10 THz range in the early 1970s [1], the first (arguably) published THz medical imaging result did not appear in the literature until the late 1990s [2]. In this work Mittleman et al. used a THz time domain system to image a burn on chicken breast skin induced by a high powered, argon ion laser. Changes were measured in the complex dielectric constant of the tissue due to both reductions in water concentration and perturbations in biochemical composition. Since the late 1990s THz medical imaging has been proposed and applied to a variety of medical imaging applications including skin [3]–[11] and breast [12]–[15] cancer margin detection, burn wound imaging [2], [16]–[19], skin hydration monitoring [20], [21], and corneal hydration measurement [22]–[24]. A check of the published dates of the preceding references indicates that THz medical imaging is essentially completing its first decade as a research discipline. The last 10 years of research have resulted in a wealth of medical THz data and results, and has helped establish initial acceptance in the medical community. Due to the presence of water in physiological tissue and the high THz absorption of water, reflective THz imaging has distinct advantages over earlier transmission-based systems, especially in in vivo applications. Furthermore the dielectric properties of water these frequencies yield easily detectable changes in THz reflectivity for small changes in hydration making water content variation an effective contrast mechanism. Variations in dielectric function have been measured in different tissue types and between cancerous and healthy tissues, and these are due largely to changes in water content [25]. These advantages coupled with the low, non-ionizing THz photon energy (0.4–40 meV) may make THz an ideal tool for in-vivo imaging of skin burns [26], [27], melanoma/carcinoma [28]–[31], corneal pathologies, and cancers. This paper will first provide an overview of THz medical imaging including a brief discussion of medical funding, a condensed history of the field, and applications currently under investigation. The review will focus on imaging specifically; for THz biomedical spectroscopy and sensing the interested reader is encouraged to access one of the following excellent review articles [12], [32]–[37]. Following the review the THz electromagnetic properties of water and its strong dependence on frequency are elucidated. The large real and imaginary components of the permittivity of water in the THz region and its prevalence in physiological tissues significantly affect THz medical imaging and water is often the dominant contrast mechanism in THz medical imagery. In Section IV, a detailed discussion of center illumination frequency and associated effects on THz imaging are provided. The THz regime covers nearly 2 decades of bandwidth and illumination frequency can significantly affect the expected spatial resolution, scattering performance, and hydration sensitivity of a THz medical imaging system. Tradeoffs are analyzed and an optimal band, covering 400–700 GHz is identified. A THz medical imaging system operating at a center frequency of 525 GHz with 125 GHz of response normalized bandwidth is presented in Section V. The system operates in reflection mode, uses a photoconductive source and Schottky diode detector, and was designed using the phenomenology arguments detailed in Section IV. Spatial resolutions of 1 mm and hydration sensitivities of 0.4% by volume were achieved. This system was used to acquire images of both ex vivo and in vivo skin burns, as well as corneal phantoms and ex vivo porcine cornea specimens.

### *A.* THz radiation and spectrum

Terahertz (THz) radiation, which occupies a large portion of the electromagnetic spectrum between the infrared and microwave bands, offers innovative image and sensing technologies that can provide information, which is not available through conventional methods (i.e. microwave and X-ray techniques.) Recently, governmental supported THz wave related fundamental research in science and application emphasized technology development has increased substantially. As THz wave (T-ray) technology improves, we believe new T-ray capabilities will impact a range of interdisciplinary fields, including: communications, imaging, medical diagnosis, health monitoring, environmental control, and chemical and biological identification. This is particularly crucial for identifying terrorist threats in homeland security (three to five years), and medical diagnosis or even clinical treatment in biomedical applications (five to ten years). T-rays offer the opportunity for transformational advances in defense and security. Recent work in our laboratory, for example, shows that T-rays have promise as a means of examining an unidentified organic powder inside an unopened paper, cardboard, or plastic container. We also are looking at T-ray spectroscopy as a method of identifying explosive compounds. Unique features in the THz spectra of these materials have been identified. A THz wave can easily penetrate and inspect the insides of most dielectric materials, which are opaque to visible light and low contrast to X-rays, making T-rays a useful complementary imaging source in this context. In addition, we have demonstrated the outstanding sensitivity of our T-ray detection systems, which can measure monolayers of certain compounds, including water. T-rays have several advantages over other sensing and imaging techniques. While microwave and X-ray imaging modalities produce density pictures, T-ray imaging also provides spectroscopic information within the THz frequency range. The unique rotational and vibrational responses of biological materials within the THz range provide information that is generally absent in optical, X-ray and NMR images. Examples of such applications to the recognition of terrorist threats include the utilization of terahertz spectroscopy in the identification of biomaterial, which has fingerprints in the terahertz range, and remote sensing and imaging of explosive targets. I will also report how THz wave imaging contributes to NASA programs in the detection of defects Dark-field imaging deals with contrast enhancement by the detection of that part of the radiation which is deflected out of the beam-propagation direction by either diffraction or scattering in the sample, and requires blocking of the ballistic part of the radiation.

### Recursive Domain equations

One of the great successes of category theory in computer science has been the development of a "unified theory" of the constructions underlying denotational semantics. In the untyped $\lambda$-calculus, any term may appear in the function position of an application. This means that a model D of the $\lambda$-calculus must have the property that given a term $t$ whose interpretation is $d \in D,$ Also, the interpretation of a functional abstraction like $\lambda x\,.\,x$ is most conveniently defined as a function from $D\,to\,D$, which must then be regarded as an element of D.

Let $\psi:[D \to D]\to D$ be the function that picks out elements of D to represent elements of $[D \to D]$ and $\phi: D \to [D \to D]$ be the function that maps elements of D to functions of D. Since $\psi(f)$ is intended to represent the function $f$ as an element of D, it makes sense to require that $\phi(\psi(f)) = f,$ that is,

$$\psi\, o\, \psi = id_{[D\to D]}$$

Furthermore, we often want to view every element of D as representing some function from D to D and require that elements representing the same function be equal – that is

$$\psi(\varphi(d)) = d$$

*or*

$$\psi\, o\, \phi = id_D$$

The latter condition is called extensionality.

These conditions together imply that $\phi\, and\, \psi$ are inverses--- that is, D is isomorphic to the space of functions from D to D that can be the interpretations of functional abstractions:

$$D \cong [D \to D]$$

Let us suppose we are working with the untyped $\lambda - calculus$, we need a solution ot the equation $D \cong A + [D \to D],$ where A is some predetermined domain containing interpretations for elements of C. Each element of D corresponds to either an element of A or an element of $[D \to D],$ with a tag. This equation can be solved by finding least fixed points of the function $F(X) = A + [X \to X]$ from domains to domains --- that is, finding domains X such that

$X \cong A + [X \to X]$, and such that for any domain Y also satisfying this equation, there is an embedding of X to Y --- a pair of maps

$$X \overset{f}{\underset{f^R}{R}} Y$$

Such that

$$f^R o\, f = id_X$$

$$f\, o\, f^R \subseteq id_Y$$

Where $f \subseteq g$ means that $f$ *approximates* $g$ in some ordering representing their information content. The key shift of perspective from the domain-theoretic to the more general category-theoretic approach lies in considering *F* not as a function on domains, but as a *functor* on a category of domains. Instead of a least fixed point of the function, *F*.

1. ***Definition*** : Let K be a category and $F : K \to K$ as a functor. A fixed point of F is a pair (A,a), where A is a K-object and $a : F(A) \to A$ is an isomorphism. A prefixed point of F is a pair (A,a), where A is a K-object and a is any arrow from F(A) to A

2. ***Definition :*** An $\omega - chain$ in a category K is a diagram of the following form:

   $$\Delta = D_o \xrightarrow{f_o} D_1 \xrightarrow{f_1} D_2 \xrightarrow{f_2} .....$$

   Recall that a cocone $\mu$ of an $\omega - chain\ \Delta$ is a K-object X and a collection of K –arrows $\{\mu_i : D_i \to X \mid i \geq 0\}$ such that $\mu_i = \mu_{i+1} o\, f_i$ for all $i \geq 0$. We sometimes write $\mu : \Delta \to X$ as a reminder of the arrangement of $\mu$'s components

   Similarly, a colimit $\mu : \Delta \to X$ is a cocone with the property that if $\nu : \Delta \to X'$ is also a cocone then there exists a unique mediating arrow $k : X \to X'$ such that for all $i \geq 0,, \nu_i = k\, o\, \mu_i$. Colimits of $\omega - chains$ are sometimes referred to as $\omega - colimits$.

Dually, an $\omega^{op} - chain$ in K is a diagram of the following form:

$$\Delta = D_o \xleftarrow{f_o} D_1 \xleftarrow{f_1} D_2 \xleftarrow{f_2} .....$$

A cone $\mu : X \to \Delta$ of an $\omega^{op} - chain\ \Delta$ is a K-object X and a collection of **K**-arrows $\{\mu_i : D_i \mid i \geq 0\}$ such that for all $i \geq 0,\ \mu_i = f_i\, o\, \mu_{i+1}$. An $\omega^{op}$-limit of an $\omega^{op} - chain\ \Delta$ is a cone $\mu : X \to \Delta$ with the property that if $\nu : X' \to \Delta$ is also a cone, then there exists a unique mediating arrow $k : X' \to X$ such that for all $i \geq 0, \mu_i\, o k = \nu_i$. We write $\perp_k$ (or just $\perp$) for the distinguish initial object of ***K,*** when it has one, and $\perp \to A$ for the unique arrow from $\perp$ to each ***K***-object A. It is also convenient to write

$\Delta^- = D_1 \xrightarrow{f_1} D_2 \xrightarrow{f_2} .....$ to denote all of $\Delta$ except $D_o$ and $f_0$. By analogy, $\mu^-$ is $\{\mu_i \mid i \geq 1\}$. For the images of $\Delta$ and $\mu$ under ***F*** we write

$$F(\Delta) = F(D_o) \xrightarrow{F(f_o)} F(D_1) \xrightarrow{F(f_1)} F(D_2) \xrightarrow{F(f_2)} .....$$

and $F(\mu) = \{F(\mu_i) \mid i \geq 0\}$

We write $F^i$ for the ***i***-fold iterated composition of ***F*** – that is,

$$F^o(f) = f, F^1(f) = F(f), F^2(f) = F(F(f))$$

,etc.

With these definitions we can state that every monitonic function on a complete lattice has a least fixed point:

3. **Lemma** Let K be a category with initial object $\perp$ and let $F : K \to K$ be a functor. Define the $\omega - chain\, \Delta$ by

   $$\Delta = \perp \xrightarrow{!\perp \to F(\perp)} F(\perp) \xrightarrow{F(!\perp \to F(\perp))} F^2(\perp) \xrightarrow{F^2(!\perp \to F(\perp))} .........$$

   If both $\mu : \Delta \to D$ and $F(\mu) : F(\Delta) \to F(D)$ are colimits, then (D,d) is an intial F-algebra, where $d : F(D) \to D$ is the mediating arrow from $F(\mu)$ to the cocone $\mu^-$

THz pulses are generated at a repetition rate of 1 ldlz via optical rectification, focused onto the sample with a cone angle of 10' and detected electro- optically after transmission through the sam- 0.6 THz ple. As the radiation from the sample is collected within a much larger cone angle of 45", a large amount of radiation diffracted and scattered within the sample is detected in addition to the ballistic radiation. In order to perform dark-field imaging, the latter is blocked by a metal beam We performed measurements on a 3-mmthick archived (formalin-ked, alcohol-dehydrated, embedded-in-paraffin) canine skin tissue sample containing a mast cell tumor in brightfield and in dark-field geometry. 2176 pixels are imaged in a total measurement time of 3 hours per image. Areas of fat, skin with hairs, connective tissue and the tumor are indicated. Images of the total loss obtained from a measurement without beam stop. The images display the deflection loss (the ratio of the power deflected relative to the power transmitted through pure paraffin) for 0.6 THz and 2.0 THz. In the tumor region, the deflection loss at 2.0 THz is exceedingly small (approaching the noise level). In order to correct for the influence of absorption in the dark-field images, we define a new quantity, the deflection coefficient, as the ratio of the deflection loss to the absorption loss. This quantity is equivalent to the relative part of the power that would be deflected from a non-absorbing or infinitely thin sample. The data taken at 2.0 THz show that the tumor region is not a strong deflector quite in contrast to the boundaries between different tissue types and the area of the skin with hairs. Comparison with the data taken at 0.6 THz sugstop.' 2.0 THz I 99.999 %. Connective tissue Skin with Tumor masked by use . 34 mm CWJl (a) Optical image of the sample; (b) and (c): Total loss in transmission, (d) and (e): Loss induced by deflection; (f) and (g): Deflection coefficient. (h) Optical image of the sample overlapped with a dark mask generated at areas where deflection loss is smaller than 0.05% and the total loss is larger than 95%. The diffraction is dominant at boundaries, while scattering dominates in the region of skin with hairs. Paper demonstrates that the tumor region may be identified by combining criteria for the total loss and the deflection loss at 2 THz as specified in the figure. In summary we believe, that THz dark-field imaging, which allows enhancing of imaging contrast especially at tissue boundaries, holds promise for the clinical distinction between benign and malignant tumors, as this is usually based on differences in the structure of the tumor boundary. For in-vivo applications reflection geometry, like in Ref. [2] has to be introduced.

THz radiation straddles microwave and infrared bands (50 GHz – 10 THz), thus combining the penetrating power of lower-frequency waves and imaging capabilities of high energy infrared radiation. Since THz radiation is not absorbed by most dry, non-polar materials, it can be used for imaging internal structures. Besides its military uses, THz radiation is employed in such important applications as spectroscopy, industrial bio-medical imaging, and scattering. What also makes THz radiation attractive is its non-ionizing photon energy, which is less than 0.1 eV at 1 THz. Several conventional devices are used presently to generate THz radiation. For example, slow-wave devices require very small structures (mm or sub-mm) in size. This complicates fabrication and alignment and results in merely milliwatts of average output power. Conventional FELs and synchrotrons are bulky and very expensive to operate. We propose a new approach for generating THz radiation that is compact, dispenses with complicated structures and relies on a well-known phenomenon called the "two-stream instability." The proposed configuration involves two low energy electron beams that are merged by a dipole magnet into a single beam and interact unstably provided the velocity difference exceeds a threshold value. Although this instability is undesirable and is usually suppressed, it can also be exploited for efficient narrowband and coherent THz production. Using a small-signal analysis, the threshold velocity difference and velocity difference for maximum gain are calculated and derived for two electron beams in a beam pipe. The calculations show an excellent agreement with a 1-D simulation of two overlapping electron beams that fill a beam pipe while interacting at 100 GHz. Preliminary 2-D PIC simulation results appear to be very promising and agree well with the theory. More 2-D PIC simulations are under way (to be followed by 3-D simulations) in order to further test and validate the proposed configuration and underlying theory.

## II. METHODOLOGY

### *A.* Reflective THz imaging power spectrum

The choice of illumination spectral density and/or detection spectral responsively for a THz imaging system can greatly affect resolution, sensitivity, scattering, and other aspects that contribute to overall image quality. Thanks to the large instantaneous bandwidth of photoconductive sources and broad spectral range of available waveguide mounted sources such as frequency multiplier chains and backward wave oscillators (BWO), researchers have nearly two decades of bandwidth to draw from corresponding to fractional bandwidths of 200%. The following sections detail some of the basic bandwidth tradeoffs for THz medical imaging. Each plot is accompanied by a highlighted region denoting the 400–700 GHz band. Center illumination frequencies in this band have been determined optimal for THz medical imaging using the following analysis and these results have influenced the design of the system.

If $X^{\alpha} \in a$ and $X^{\beta} \in k\left[X_1, \ldots X_n\right]$ , then $X^{\alpha} X^{\beta} = X^{\alpha+\beta} \in a$ , and so $A$ satisfies the condition $(*)$. Conversely,

$$(\sum_{\alpha\in A} c_\alpha X^\alpha)(\sum_{\beta\in\mathbb{N}^n} d_\beta X^\beta) = \sum_{\alpha,\beta} c_\alpha d_\beta X^{\alpha+\beta}$$
$(finite\ sums)$,

and so if $A$ satisfies $(*)$, then the subspace generated by the monomials $X^\alpha, \alpha\in a$, is an ideal.

The proposition gives a classification of the monomial ideals in $k[X_1,...X_n]$: they are in one to one correspondence with the subsets $A$ of $\mathbb{N}^n$ satisfying $(*)$. For example, the monomial ideals in $k[X]$ are exactly the ideals $(X^n)$, $n\ge 1$, and the zero ideal (corresponding to the empty set $A$). We write $\langle X^\alpha \mid \alpha\in A\rangle$ for the ideal corresponding to $A$ (subspace generated by the $X^\alpha, \alpha\in a$).

LEMMA 0.4. Let $S$ be a subset of $\mathbb{N}^n$. The the ideal $a$ generated by $X^\alpha, \alpha\in S$ is the monomial ideal corresponding to

$$A \overset{df}{=} \{\beta\in\mathbb{N}^n \mid \beta-\alpha\in\mathbb{N}^n,\quad some\ \alpha\in S\}$$

Thus, a monomial is in $a$ if and only if it is divisible by one of the $X^\alpha, \alpha\in| S$

PROOF. Clearly $A$ satisfies $(*)$, and $a\subset\langle X^\beta \mid \beta\in A\rangle$. Conversely, if $\beta\in A$, then $\beta-\alpha\in\mathbb{N}^n$ for some $\alpha\in S$, and $X^\beta = X^\alpha X^{\beta-\alpha}\in a$. The last statement follows from the fact that $X^\alpha \mid X^\beta \Leftrightarrow \beta-\alpha\in\mathbb{N}^n$.

Let $A\subset\mathbb{N}^n$ satisfy $(*)$. From the geometry of $A$, it is clear that there is a finite set of elements $S=\{\alpha_1,...\alpha_s\}$ of $A$ such that

$$A=\{\beta\in\mathbb{N}^n \mid \beta-\alpha_i\in\mathbb{N}^2,\ some\ \alpha_i\in S\}$$

(The $\alpha_i$'s are the corners of $A$) Moreover, $a \overset{df}{=} \langle X^\alpha \mid \alpha\in A\rangle$ is generated by the monomials $X^{\alpha_i}, \alpha_i\in S$.

DEFINITION 0.3. For a nonzero ideal $a$ in $k[X_1,...,X_n]$, we let $(LT(a))$ be the ideal generated by

$$\{LT(f) \mid f\in a\}$$

LEMMA 0.8 Let $a$ be a nonzero ideal in $k[X_1,...,X_n]$; then $(LT(a))$ is a monomial ideal, and it equals $(LT(g_1),...,LT(g_n))$ for some $g_1,...,g_n\in a$.

PROOF. Since $(LT(a))$ can also be described as the ideal generated by the leading monomials (rather than the leading terms) of elements of $a$.

THEOREM 0.11. Every *ideal* $a$ in $k[X_1,...,X_n]$ is finitely generated; more precisely, $a=(g_1,...,g_s)$ where $g_1,...,g_s$ are any elements of $a$ whose leading terms generate $LT(a)$

PROOF. Let $f\in a$. On applying the division algorithm, we find

$$f = a_1 g_1 + ... + a_s g_s + r, \qquad a_i, r\in k[X_1,...,X_n]$$

, where either $r=0$ or no monomial occurring in it is divisible by any $LT(g_i)$. But $r = f - \sum a_i g_i \in a$, and therefore $LT(r)\in LT(a) = (LT(g_1),...,LT(g_s))$, implies that every monomial occurring in $r$ is divisible by one in $LT(g_i)$. Thus $r=0$, and $g\in(g_1,...,g_s)$.

DEFINITION 0.11. A finite subset $S=\{g_1,|...,g_s\}$ of an ideal $a$ is a standard ($(Gr\ddot{o}bner)$ bases for $a$ if $(LT(g_1),...,LT(g_s)) = LT(a)$. In other words, S is a standard basis if the leading term of every element of $a$ is divisible by at least one of the leading terms of the $g_i$.

**THEOREM 1.1** (Hilbert Basis Theorem). *The ring $k[X_1,...,X_n]$ is Noetherian i.e., every ideal is finitely generated.*

PROOF. For $n=1$, $k[X]$ is a principal ideal domain, which means that every ideal is generated by single element. We shall prove the theorem by induction on $n$. Note that the obvious map

$$k[X_1,...X_{n-1}][X_n] \to k[X_1,...X_n]$$

is an isomorphism – this simply says that every polynomial $f$ in $n$ variables $X_1,...X_n$ can be expressed uniquely as a polynomial in $X_n$ with coefficients in $k[X_1,..., X_n]$:

$$f(X_1,...X_n) = a_0(X_1,...X_{n-1})X_n^r + ... + a_r(X_1,...X_{n-1})$$

Thus the next lemma will complete the proof

LEMMA 1.3. If $A$ is Noetherian, then so also is $A[X]$

PROOF. For a polynomial

$$f(X) = a_0X^r + a_1X^{r-1} + ... + a_r, \qquad a_i \in A, \qquad a_0 \neq 0,$$

$r$ is called the degree of $f$, and $a_0$ is its leading coefficient. We call 0 the leading coefficient of the polynomial 0.

Let $a$ be an ideal in $A[X]$. The leading coefficients of the polynomials in $a$ form an ideal $a'$ in $A$, and since $A$ is Noetherian, $a'$ will be finitely generated. Let $g_1,..., g_m$ be elements of $a$ whose leading coefficients generate $a'$, and let $r$ be the maximum degree of $g_i$.

Now let $f \in a$, and suppose $f$ has degree $s > r$, say, $f = aX^s + ...$ Then $a \in a'$, and so we can write

$$a = \sum b_i a_i, \qquad b_i \in A,$$

$$a_i = \text{leading coefficient of } g_i$$

Now

$$f - \sum b_i g_i X^{s-r_i}, \quad r_i = \deg(g_i),$$

has degree $< \deg(f)$. By continuing in this way, we find that

$$f \equiv f_t \qquad \mathrm{mod}(g_1,...g_m)$$

With $f_t$ a polynomial of degree $t < r$

For each $d < r$, let $a_d$ be the subset of $A$ consisting of 0 and the leading coefficients of all polynomials in $a$ of degree $d$; it is again an ideal in $A$. Let $g_{d,1},..., g_{d,m_d}$ be polynomials of degree $d$ whose leading coefficients generate $a_d$. Then the same argument as above shows that any polynomial $f_d$ in $a$ of degree $d$ can be written

$$f_d \equiv f_{d-1} \qquad \mathrm{mod}(g_{d,1},...g_{d,m_d})$$

With $f_{d-1}$ of degree $\leq d-1$. On applying this remark repeatedly we find that

$$f_t \in (g_{r-1,1},...g_{r-1,m_{r-1}},...g_{0,1},...g_{0,m_0})$$

Hence

$$f_t \in (g_1,...g_m g_{r-1,1},...g_{r-1,m_{r-1}},..., g_{0,1},..., g_{0,m_0})$$

and so the polynomials $g_1,..., g_{0,m_0}$ generate $a$

*B.* **Hydration Sensitivity and Penetration Depth**

For many THz medical imaging applications, measurement sensitivity is determined by quantifying the expected change in THz tissue reflectivity for a given change in water volume fraction at a particular volume fraction. As discussed above, recent literature has demonstrated through both theory and experiment that dielectric models of water can accurately describe THz interaction with biological material, and much of the contrast in tissue imaging is primarily due to water concentration gradients. Basic calculations on hydration sensitivity can be explored using the dielectric properties of water and simple Fresnel equations. Although illumination beams in THz imaging systems are generally Gaussian [100], the Rayleigh length of the beam after focusing is generally longer than the depth of penetration/interaction of the THz beam in tissue, given the large attenuation constant. Accurate results can therefore be obtained using plane wave analysis. (For more in-depth calculations see the k-space method [101]). For this study we model the tissue as a homogenous mixture of water and biological background which is constructed as a lossless dielectric with an index of 2. The reflection coefficient of a half space of tissue in air

**Definition.** A fibering or fiberation $p: X \to B$ is a proper, flat morphism of finite presentation. A fibration of curves of genus $g$ is a fibration $p: X \to B$ such that every fibre $X_b$ is a curve of genus $g$.

**Definition.** An elliptic fibration $p: X \to B$ is a fibration of genus one curves together with a section $e: B \to X$ of $p$ such that the geometric fibres are irreducible reduced curves. A polarized fibration of $K3$ surfaces is $K3$ fibration $p: X \to B$ is a fibration of surfaces such that the geometric fibres of $p$ are $K3$ surfaces together with a class $\xi \in Pic(X)$ such that its restriction to each fibre $\xi_b \in Pic(X_b)$ is the class of an ample line bundle.

It is possible that all fibres over geometric points are singular rational curves of genus 1 in a genus one fibration over a curve. This happens only in characteristic $p = 2$ and 3, but in general, only a finite number of fibres are singular.

**General Fibre.** If $f: X \to B$ is a fibration with general fibre $F$, then except for a finite set of $b \in B$, the fibre $X_b$ of $f$ over $b$ is an irreducible curve. We can apply the genus formula to obtain the genus $q(F)$ as $2q(F)-2 = K_x.F$ since $F.F = 0$. We have an elliptic fibration if and only if $K_X.F = 0$. In the case of a surface over a curve, the fibres are curves or more generally one dimensional schemes of the form $D = \sum_i n_i E_i$ where the $E_i$ are irreducible curves. Hence the fibre is this advisor $D$.

**Proposition.** Let $V$ be an inner product space over $Q$ generated by vectors $e_i$ for $i \in I$ with $(e_i \mid e_j) \geq 0$ for all $i \neq j$. If there exists a vector $z = \sum_i a_i e_i$ with all $a_i > 0$ such that $(z \mid e_j) = 0$ for all $j$, then we have $(z \mid x) = 0$ and $(x \mid x) \leq 0$ for all $x \in V$.

*Proof.* Since any $x$ is a linear combination of the $e_j$ and $(z \mid e_j) = 0$ for all indices $j$, it follows that $(z \mid x) = 0$. for the negative definite statement, we write any $x \in V$ as $x = \sum_i c_i a_i e_i$ where $c_i \in Q$ and calculate

$$(x \mid x) - \sum_i c_i^2 a_i^2 (e_i \mid e_i) =$$

$$\sum_{i \neq j} c_i c_j (a_i e_i \mid a_j e_j)$$

$$\leq \sum_{i \neq j} \frac{1}{2}(c_i^2 + c_i^2)(a_i e_i \mid a_j e_j)$$

$$= \sum_i \frac{1}{2} c_i^2 (a_i e_i \mid z - a_i e_i)$$

$$+ \sum_j \frac{1}{2} c_j^2 (a_j e_j \mid z - a_j e_j)$$

$$= -\sum_i c_i^2 a_i^2 (e_i \mid e_i)$$

Thus the sequence of inequalities gives $(x \mid x) \leq 0$. This proves the proposition.

**Theorem.** If an effective divisor $D$ on a $K3$ surface $X$ satisfies the conditions $(D^2) = 0$ and $D.C > 0$ for every curve $C$ on $X$, then the linear system $|D|$ contains a divisor of the form $mE$ where $m > 0$ and $E$ is an elliptic curve.

*Proof.* We must show that $|D|$ contains a divisor with only one component.

Step 1. Let $D' \in |D|$, and consider decompositions $D' = \sum_{i=1}^r a_i C_i$ where the $C_i$ are distinct and irreducible and $a_i > 0$. Assume there are at least two indices $i$. Then we have $D.C \geq 0$ and $C_i.C_j \geq 0$ for $i \neq j$, and this implies that $C_i^2 = 0$. We show that the self intersection $C_i^2 = 0$ for some $D' \in |D|$. For this consider an embedding $X \to P^N$ with general hyperplane $H \subset P^N$ and hyperplane section $X \cap H$ of $X$. In the intersection $D.H = D'.H = \sum_{i=1}^r a_i (C_i.H)$ the terms $C_i.H$ are just the degrees of the embeddings $C_i \to P^N$. Hence the positive integers $r$ and $a_i$ are bounded. Therefore, the number of $D' \in |D|$ with all $C_i.C_i < 0$ is finite in number, while the linear system $|D|$ is infinite.

We have shown the existence of $D' \in |D|$ which decomposes as a sum of the form $D' = mE + D'' = mE + \sum_{i=1}^r a_i C_i$ where $E.E = 0$ and $m \geq 0$.

Step 2. In the decompositions $D' = mE + D''$ as in step 1, we have two further intersection properties: $E.D'' = 0$ and $D''.D'' = 0$. For we calculate $0 = D'.D' = mE.D' + D''.D'$, we have $E.D' = 0$ so that $D''.D' = 0$. Moreover, $E.D'' \geq 0$ since they have no common components, and from the following two relations $0 = D'.D'' = mE.D'' + D''.D''$ and $0 = D'.D' = 2mE.D'' + D''.D''$ we see that $E.D'' = 0$ and $D''.D'' = 0$.

Step 3. Among the $D' \in |D|$ decomposed as $D' = mE + D''$ with $E.E = 0$, so that $E.D'' = 0$, $D''.D'' = 0$ by step 2, we choose the one with $E''.H$ minimal as a natural number for $H$ a hyperplane section. It remains to show that $D'' = 0$ to complete the proof, and this we do by deriving a contradiction assuming $D'' \neq 0$. If $D'' \neq 0$, then $l(D) \geq 2$, and we can apply the above considerations in steps 1 and 2 for $|D|$ to

$\left|D''\right|$. There exists $D(1)\in\left|D''\right|$ with the decomposition

$D''$ linearly equivalent to $D(1)=nE+D(2)$

So that $D'$ is equivalent to $(m+n)E+D(2)$. This means that intersecting this $D(2)$ with the hyperplane $H$ we have

$$H.D(2)=H.D(1)-H.(mE)=H.D'-H.(mE)<H.D''$$

Which contradicts the minimal character of $K''$. Hence $D''=0$ for the minimal case and thus $D'=mH$. This proves the theorem.

**Fibrations of 3 Dimensional**

For a 3-fold, we can look for a fibration by either surfaces or curves. In the first case, we would start with a divisor or line bundle with the intersection properties of a fibre, and in the second case we would start with a divisor with the selfintersection properties of a fibre of curves. The divisors in this picture would be numerically effective divisors.

**Definition.** A divisor $D$ on a variety is numerically effective, or nef, provided $D.C\geq 0$ for all curves $C$ on $X$. A line bundle $L$ is numerically effective provided it is of the form $L=O(D)$ where $D$ is a numerically effective divisor.

*C.* **Scattering**

Scattering from rough surfaces is a well known and often observed problem in optics, and has been studied in the THz band for simple cases. In THz medical imaging, particularly the imaging of skin, typical target feature sizes approach hundreds of micrometers, placing them directly in the middle of the wavelength bands of interest. This poses a significant problem for hydration sensing, where small changes in hydration dependent reflectivity may be masked by random scattering caused by target geometry. This aspect has led many researchers to employ a window composed of a lossless dielectric to flatten the field during THz medical imaging experiments. However windows add additional system complexity as well as raise issues of sterilization and should be avoided when possible. A common method used to model frequency dependent scattering in the THz regime is the Rayleigh roughness factor where is the standard deviation of the surface roughness, the illumination angle of incidence, and the illumination wavelength. The Rayleigh roughness factor describes the average fraction of power reflected in the specular direction for plane wave illumination and assumes a Gaussian probability density function (pdf) of surface profile heights. Equation (8) was simulated for standard deviations of 30, 70, and 150 m and incidence angles of 0 , 30 , and 45 , thus encompassing typical tissue surface roughnesses and common illumination angles. The results are shown in Fig. 4. As expected, lower frequencies are much more robust to scattering than higher frequencies, and tissues appear more specular in the millimeter wave range than they do in the sub-millimeter. An interesting consequence of the Rayleigh roughness factor is the effect of incidence angle as displayed. Oblique illumination reduces scatter in the non-specular direction and in certain circumstances may motivate operating at glancing incidence at the expense of increased spot size due to beam smearing. Spot size also significantly affects the scattering performance of the imaging system. The expected variance in return signal of a Gaussian beam swept across a random rough surface is difficult to model and as such closed form expressions describing the statistics are difficult to produce [7]. In lieu of a mathematical model, we present images of ex vivo porcine skin acquired with two pairs of off-axis parabolic (OAP) mirrors incorporating differing incidence angles and spot sizes.

**Lemma 1.1.1** The metric space $(P((\Omega,B)),d_G)$ of all probability measures on a measurable space $(\Omega,B)$ with a countably generated sigma-field is separable if $G$ contains a countable generating field. If also $G$ is standard (as is possible if the underlying measurable space $(\Omega,B)$ is standard), then also $(P,d_G)$ is complete and hence Polish.

*Proof*: For each $n$ let $A_n$ denote the set of nonempty intersection sets or atoms of $\{F,\ldots,F_n\}$, the first $n$ sets in $G$. (These are $\Omega$ sets, events in the original space) For each set $G\in A_n$ choose an arbitrary point $x_G$ such that $x_G\in G$. We will show that the class of all measures of the form

$$r(F)=\sum_{G\in A_n}P_G 1_F(x_G),$$

Where the $p_G$ are nonnegative and rational and satisfy

$$\sum_{G\in A}p_G=1,$$

Forms a dense set in $P((\Omega,B))$. Since this class is countable, $P((\Omega,B))$ is separable. Observe that we are approximating all measures by finite sums of point masses. Fix a measure $m\in(P((\Omega,B)),d_G)$ and an $\varepsilon>0$. Choose $n$ so large that $2^{-n}<\varepsilon/2$. Thus to match up two measures in $d=d_G$, implies that we must match up the probabilities of the first

$n$ sets in $G$ since the contribution of the remaining terms is less that $2^{-n}$. Define

$$r_n(F) = \sum_{G \in A_n} m(G) 1_F(x_G)$$

And note that

$$m(F) = \sum_{G \in A_n} m(G) m(F \mid G),$$

Where $m(F \mid G = m(F \cap G)/m(G)$ is the elementary conditional probability of $F$ given $G$ if $m(G) > 0$ and is arbitrary otherwise. For convenience we now consider the preceding sums to be confined to those $G$ for which $m(G) > 0$.

Since the $G$ are the atoms of the first $n$ sets $\{F_i\}$, for any of these $F_i$ either $G \subset F_i$ and hence $G \cap F_i = G$ or $G \cap F_i = \phi$. In the first case $1_{F_i}(x_G) = m(F_i \mid G) = 1$, and in the second case $1_{F_i}(x_G) = m(F_i \mid G) = 0$, and hence in both cases

$$r_n(F_i) = m(F_i); i = 1, 2, ..., n$$

This implies that

$$d(r_n, m) \le \sum_{i=n+1}^{\infty} 2^{-i} = 2^{-n} \le \frac{\varepsilon}{2}$$

Enumerate the atoms of $A_n$ as $\{G_l; l = 1, 2, ..., L$, Where $L \le 2^n$. For all $l$ but the last $(l = L)$ pick a rational number $p_{G_l}$ such that

$$m(G_l) - 2^{-l}\frac{\varepsilon}{4} \le p_{G_l} \le m(G_l);$$

That is, we choose a rational approximation to $m(G_l)$ that is slightly less than $m(G_l)$. We define $p_{G_l}$ to force the rational approximations to sum to 1:

$$p_{G_l} = 1 - \sum_{l=1}^{L-1} p_{G_l}$$

Clearly $p_{G_l}$ is also rational and

$$\left| p_{G_l} - m(G_l) \right| = \left| \sum_{l=1}^{L-1} p_{G_l} - \sum_{l=1}^{L-1} m(G_l) \right|$$
$$\le \sum_{l=1}^{L-1} \left| p_{G_l} - m(G_l) \right|$$
$$\le \frac{\varepsilon}{4} \sum_{l=1}^{L-1} 2^{-l} \le \frac{\varepsilon}{4}$$

Thus

$$\sum_{G \in A_n} \left| p_{G_l} - m(G) \right| = \sum_{l=1}^{L-1} \left| p_{G_l} - m(G_l) \right| + \left| p_{G_l} - m(G_l) \right| \le \frac{\varepsilon}{2}$$

Define now the measure $t_n$ by

$$t_n(F) = \sum_{G \in A_n} p_G 1_F(x_G)$$

And observe that

$$d(t_n, r_n) = \sum_{i=1}^{n} 2^{-i} \left| t_n(F_i) - r_n(F_i) \right|$$
$$\le \sum_{i=1}^{n} 2^{-i} \left| \sum_{G \in A_n} 1_{F_i}(p_G - m(G)) \right|$$
$$\le \sum_{i=1}^{n} 2^{-i} \sum_{G \in A_n} \left| p_G - m(G) \right| \le \frac{\varepsilon}{4}$$

We now know that

$$d(t_n, m) \le d(t_n, r_n) + d(r_n, m) \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

Which completes the proof of separabillity.

Next assume that $m_n$ is a Cauchy sequence of measures. From the definition of $d$ this can only be if also $m_n(F_i)$ a Cauchy sequence of real numbers is for each $i$. Since the real line is complete, this means that for each $im_n(F_i)$ converges to something, say $\alpha(F_i)$. The set function $\alpha$ is defined on the class $G$ and is clearly nonnegative, normalized, and finitely additive. Hence if $G$ is also standard, then $\alpha$ extends to a probability measure on $(\Omega, B)$; that is, $\alpha \in P((\Omega, B))$. By construction $d(m_n, \alpha) \to 0$ as $n \to \infty$, and hence in this case $P((\Omega, B))$ is complete.

**Lemma 1.1.2** If $f$ is a nonnegative continuous function and $(\Omega, B)$ and $d_G$ are as previously, then $d_G(m_n, m) \to 0$ implies that $\limsup_{n \to \infty} E_m f \le E_m f$

Proof: For any $n$ we can divide up $[0, \infty)$ into the countable collection of intervals $[k/n, (k+1)/n)$, $k = 0, 1, 2, ...$ that partition the nonnegative real line. Define the sets $G_k(n) = \{r : k/n \le f(r) < (k+1)/n\}$ and define the closed sets $F_k(n) = \{r : k/n \le f(r)\}$. Observe that $G_k(n) = F_k(n) - F_{k+1}(n)$. Since $f$ is nonnegative for any $n$

$$\sum_{k=0}^{\infty} \frac{k}{n} m(G_k(n)) \le E_m f \sum_{k=0}^{\infty} \frac{k+1}{n} m(G_k(n)) = \sum_{k=0}^{\infty} \frac{k}{n} m(G_k(n)) + \frac{1}{n}$$

The sum can be written as

$$\frac{1}{n}\sum_{k=0}^{\infty} k(m(F_k(n)) - m(F_{k+1}(n)) = \frac{1}{n}\sum_{k=0}^{\infty} m(F_k(n)),$$

and therefore

$$\frac{1}{n}\sum_{k=0}^{\infty} m(F_k(n)) \le E_m f$$

By a similar construction

$$E_{m_n} f < \frac{1}{n}\sum_{k=0}^{\infty} m_n(F_k(n)) + \frac{1}{n}$$

And hence from the property for closed sets

$$\lim_{n\to\infty} \sup E_{m_n} f \le \frac{1}{n}\sum_{k=0}^{\infty} \lim \sup m_n(F_k(n)) + \frac{1}{n}$$

$$\le \frac{1}{n}\sum_{k=0}^{\infty} m(F_k(n)) + \frac{1}{n} \le E_m f + \frac{1}{n}$$

Since this is true for all $n$, the lemma is proved.

**Theorem 1.0** Suppose that $m_X$ and $m_Y$ are stationary process distributions with a common standard alphabet $A$ and that $\rho_1$ is a pseudo-metric on $A$ and that $\rho_n$ is defined on $A^n$ in an additive fashion. Then

(a) $\lim_{n\to\infty} n^{-1}\overline{\rho}_n(P_{X^n}, P_{Y^n})$ exists and equals $\sup_n n^{-1}\overline{\rho}_n(P_{X^n}, P_{Y^n})$.

(b) $\overline{\rho}_n$ and $\overline{\rho}$ are pseudo-metrics. If $\rho_1$ is a metric, then $\overline{\rho}_n$ and $\overline{\rho}$ are metrics.

(c) If $m_X$ and $m_Y$ are both $IID$, then $\overline{\rho}(m_X, m_Y) = \overline{\rho}_1(P_{X_0}, P_{Y_0})$.

(d) Let $P_s = P_s(m_X, m_Y)$ denote the collection of all stationary distributions $p_{XY}$ having $m_X$ and $m_Y$ as marginals, that is, distributions on $\{X_n, Y_n\}$ with coordinate processes $\{X_n\}$ and $\{Y_n\}$ having the given distributions. Define the process distance measure $\overline{\rho}'$

$$\overline{\rho}'(m_X, m_Y) = \inf_{pXY\in P_s} E_{pXY}\rho(X_0, Y_0)$$

Then

$$\overline{\rho}(m_X, m_Y) = \overline{\rho}'(m_X, m_Y);$$

That is the limit of the finite dimensional minimizations is given by a minimization over stationary processes.

(e) Suppose that $m_X$ and $m_Y$ are both stationary and ergodic. Define $P_e = P_e(m_X, m_Y)$ as the subset of $P_s$ containing only ergodic processes, then

$$\overline{\rho}(m_X, m_Y) = \inf_{pXY\in P_e} E_{pXY}\rho(X_0, Y_0),$$

(f) Suppose that $m_x$ and $m_Y$ are both stationary and ergodic. Let $G_X$ denote a collection of generic sequences for $m_X$. Recall that by measuring relative frequencies on generic sequences one can deduce the underlying stationary and ergodic measure that produced the sequence. Similarly let $G_Y$ denote a set of generic sequences for $m_Y$. Define the process distance measure

$$\overline{\rho}''(m_X, m_Y) = \inf_{x\in G_X, y\in G_Y} \lim_{n\to\infty}\sup \frac{1}{n}\sum_{i=0}^{n-1} \rho_1(x_0, y_0)$$

Then

$$\overline{\rho}(m_X, m_Y) = \overline{\rho}''(m_X, m_Y)$$

That is, the $\overline{\rho}$ distance gives the minimum long term time average distance obtainable between generic sequences from two processes.

(g) The infima defining $\overline{\rho}_n$ and $\overline{\rho}'$ are actually minima.

Proof: (a) Suppose that $p^N$ is a joint distribution on $(A^N \times A^N, B_A^N \times B_A^N)$ describing $(X^N, Y^N)$ that approximately achieves $\overline{\rho}_N$, e.g., $\rho^N$ has $P_{X^N}$ and $P_{Y^N}$ as marginals and for $\varepsilon > 0$

$$E_{p^N}\rho(X^N, Y^N) \le \overline{\rho}_N(P_{X^N}, P_{Y^N}) + \varepsilon$$

For any $n < N$ let $p^n$ be the induced distribution for $(X^N, Y^N)$ and $p_n^{N-n}$ that for $(X_n^{N-n}, Y_n^{N-n})$

Then since the processes are stationary $p^n \in P_n$ and $p_n^{N-n} \in P_{N-n}$ and hence

$$E_{p^N}\rho(X^N, Y^N)$$
$$= E_{p^n}\rho(X^n, Y^n)$$
$$+ E_{p_n^{N-n}}\rho(X_n^{N-n}, Y_n^{N-n})$$
$$\geq \overline{\rho_n}(P_{X^n}, P_{Y^n}) + \overline{\rho}_{N-n}(P_{X^{N-n}}, P_{Y^{N-n}})$$

Since $\varepsilon$ is arbitrary we have shown that if $a_n = \overline{\rho_n}(P_{X^n}, P_{Y^n})$, then for all $N > n$

$$a_N \geq a_n + a_{N-n};$$

That is, the sequence $a_n$ is superadditive. Since the negative of a superadditive sequence is a subadditive sequence

$$\lim_{n\to\infty}\frac{a_n}{n} = \sup_{n\geq 1}\frac{a_n}{n},$$

Which proves (a).

(b) Since $\rho_1$ is a pseudo-metric, $\rho_n$ is also a pseudo-metric and hence $\overline{\rho_n}$ is nonnegative and symmetric in its arguments. Thus $\overline{\rho}$ is also nonnegative and symmetric in its arguments. Thus we need only prove that the triangle inequality holds. Suppose that $m_X, m_Y$ and $m_Z$ are three process distributions on the same sequence space. Suppose that $p$ approximately yields $\overline{\rho_n}(P_{X^n}, P_{Z^n})$ and that $r$ approximately yields $\overline{\rho_n}(P_{Z^n}, P_{Y^n})$. These two distributions can be used to construct a joint distribution $q$ for the triple $(X^n, Z^n, Y^n)$ such that $X^n \to Z^n \to Y^n$ is a Markov chain and such that the coordinate vectors have the correct distribution. In particular, $r$ and $m_Z^n$ together imply a conditional distribution $p_{Y^n} | Z^n$ and $p$ and $P_{X^n}$ together imply a conditional distribution $p_{Z^n} | X^n$. Since the spaces are standard these can be chosen to be regular conditional probabilities and hence the distribution specified by its values on rectangles as a cascade of finite dimensional channels: If we define $v_x(F) = P_{Z^n|X^n}(F | x^n)$, then

$$q(F\times G\times F) = \int_D dP_{X^n}(x^n)\int_G dv_x(z^n)P_{Y^n|Z^n}(F | z^n)$$

For this distribution, since $\rho_n$ is a pseudo-metric,

$$\overline{\rho_n}(P_{X^n}, P_{Y^n}) \leq E\rho_n(X^n, Y^n)$$
$$\leq E\rho_n(X^n, Y^n)$$

## III. REFLECTIVE THZ MEDICAL IMAGING SYSTEM

A block diagram of the pulsed THz imaging system [11], [18], used to generate the images and a CAD drawing of the system is displayed. The THz source was a 9 m 9 m photoconductive switch [40], [45] pumped by a 780 nm femtosecond (fs) laser with a 230 fs pulse width, 20 MHz repetition rate, and 8 mW of average power. At high DC-bias fields (200 V/9 m gap 222 kV/cm) the source produced an optical to quasioptical (THz) conversion efficiency of 1% yielding average powers of up to 46 uWacross 1 THz of bandwidth [45]. The switch was mounted on the backside of a high resistivity silicon hyperhemisphere and positioned 60 mm away from a 76.2 mm effective focal length (EFL), 25.4 mm clear aperture OAP mirror as this numerical aperture was found to be the best match to the photo conductive switch beam pattern. The collimated beam was directed towards a THz OAP objectivewhere it was focused onto the target at a 30 , 14 , or 9 angle for 25.4 mm, 50.8 mm, and 76.2 mm EFL OAP pairs respectively. The reflected beam was collimated by a third parabolic mirror and then focused using a 25.4 mm (EFL) OAP into the feed horn of a 0-bias Schottky diode [116] detector mounted in a WR1.5 waveguide. Following the THz rectifier was a gated receiver consisting of a low-noise pulse amplifier, a double-balanced mixer, and a low pass filter (integrator). The rectified THz pulse was amplified ( 40 dB, 10 GHz) and then coupled to the RF port of a double-balanced mixer. The gating was realized by driving the LO port of the mixer with a reference RF pulse generated from sampling the mode-locked laser using a free space 99/1 beam splitter, a photodiode and a broadband amplifier. The reference pulse was passed through an RF delay line adjusted so that the pulse arrives at the mixer synchronous to the amplified THz pulse. The DC voltage from the IF port of the mixer was sent through a low pass filter, amplified with an audio frequency instrumentation amplifier, and sampled using a 14 bit DAQ with a 0.8 ms time constant. Pixels are generated by raster scanning the target in the x and y directions, using stepper motors. As a consequence of the high optical to THz conversion efficiency, the system produced sufficiently high SNR with a low power femtosecond fiber laser (1560 nm mode locked EDFA PPLN SHG [117]). This allowed a compact, imaging head integrating

780 nm and THz optics. The imaging head measured 10 cm 15 cm 25 cm and comsumed a total system volume of less than 3750 cm . A major practical advantage of the receiver architecture shown schematically in Fig. 7 is the number of optical components. This design minimizes laser alignment and down-converts the THz signal to base band immediately, making the system more robust to misalignment. The compact size and robust layout improve the portability of the system, and have allowed reliable operation in the animal operating room environment.

### A. Effective Illumination Band

The effective center frequency and bandwidth of the system are constrained by the switch power spectral density (PSD) the detector spectral responsivity. Waveguide mounted, 0-bias Schottky diodes are convenient detectors in the THz regime as these devices offer high room temperature responsivity ( 1000 V/W), low NEP ( 100 pW/Hz ), and extremely broad video bandwidth (1–14 GHz). Furthermore, the waveguides are typically terminated in pyramidal feed horns with well-known and well behaved beam patterns exhibiting extremely low side lobes. Another advantage of the waveguide mounting is that it provides well-defined pass bands with a sharp cut on frequency and a relatively sharp roll-off due to the emergence of higher order modes. This is especially useful when used as an incoherent detector of broadband illumination as it ensures an unambiguous operational band. The normalized power spectral density of the photoconductive switch is displayed in Fig. 9 superimposed on the normalized Schottky diode spectral responsivity. The switch spectrum was acquired with a Fourier Transform Infrared (FTIR) spectrometer and He-cooled composite bolometer. The detector spectral responsivity was measured with a THz photomixing setup [118].

### B. Optical Characterization

Spot size and depth of focus measurements are displayed. The spot was measured using a knife edge target with the edge swept through the beam (in the x direction as defined in Figs. 7 and 8), and is defined with the standard 10–90 edge response criteria. The data follows the fit (dotted line) predicted by the 2D integration of TEM Gaussian beam and yields a 10–90 dimension of 1.1 mm. The depth of focus (DOF) was measured by translating a polished metal reflector in and out of the focal plane ( direction as defined by the axes in Fig. 8) and the OAP ELFs and measures a total of 4 mm full width at half maximum (FWHM). Traces of the reference pulse (grey dotted) and rectified THz pulse (black solid) feeding the double balanced mixer. The shaded (pink) envelope around the rectified THz is a 13 ps window corresponding to differences in pulse arrival time from target height changes of 2 mm (system DOF). Superimposed on the data is a Gaussian fit whose shape is predicted by Gaussian beam transverse mode matching [1]. A slight asymmetry about the maximum is visible and is due to unequal beam walk off as the target is moved above and below the focal plane. The DOF is limited primarily by the optics of the system and not the pulse multiplication of the receiver. The delay line was manually scanned at the extremum of the DOF sweep (6 mm) and found to have minimal effect on the synchronicity of the rectified THz pulse and reference pulse.

We start with continuity: $\frac{\partial(\rho u)}{\partial x}+\frac{\partial(\rho v)}{\partial y}=0$

Multiply this by edge velocity $u_e(x)$. Integrate by parts.

$$u_e\frac{\partial(\rho u)}{\partial x}=\frac{\partial(\rho u u_e)}{\partial x}-\rho u\frac{\partial u_e}{\partial x}$$

$$u_e\frac{\partial(\rho v)}{\partial y}=\frac{\partial(\rho v u_e)}{\partial y}$$

$$\text{Result}:\frac{\partial(\rho u u_e)}{\partial x}-\rho u\frac{\partial u_e}{\partial x}+\frac{\partial(\rho v u_e)}{\partial y}=0$$

(1)

Subtract from this the u-momentum equation:

$$\frac{\partial(\rho u^2)}{\partial x}+\frac{\partial(\rho u v)}{\partial y}+\frac{\partial p}{\partial x}=\frac{\partial \tau_{xy}}{\partial y}$$

(2)

Use also the Euler equation: □ VdV+dp=0 which gives at boundary layer edge

$$\rho_e u_e\frac{\partial u_e}{\partial x}+\frac{\partial p}{\partial x}=0$$

(3)

Result after some minor rearranging:

$$\frac{\partial}{\partial x}\left[\rho_e u_e^2\left(1-\frac{u}{u_e}\right)\frac{\rho u}{\rho_e u_e}\right]+\frac{\partial}{\partial y}(\rho v u_e-\rho v u)+\rho_e u_e\frac{\partial u_e}{\partial x}\left(1-\frac{\rho u}{\rho_e u_e}\right)=-\frac{\partial \tau_{xy}}{\partial y}$$

(4)

Integrate equation (4) from y=0 to y=□ .

$$\frac{\partial}{\partial x}\left[\rho_e u_e^2\int_0^\delta\left(1-\frac{u}{u_e}\right)\frac{\rho u}{\rho_e u_e}dy\right]+(\rho v u_e-\rho v u)\Big|_0^\delta+\rho_e u_e\frac{\partial u_e}{\partial x}\int_0^\delta\left(1-\frac{\rho u}{\rho_e u_e}\right)dy=-\int_0^\delta\frac{\partial \tau_{xy}}{\partial y}dy$$

Use definitions of displacement thickness and momentum thickness. Second term goes to zero, since v=- at wall, and u=$u_e$ at the edge. We get:

$$\frac{\partial}{\partial x}\left(\rho_e u_e^2\theta\right)+\rho_e u_e\frac{\partial u_e}{\partial x}\delta^*=\tau_{wall} \qquad (5)$$

Equation (5) works for laminar and turbulent, compressible and incompressible flows.

Expand first term. Simplify this for incompressible flows, by assuming    e is constant.
We use ordinary derivatives from now on, since there are no derivatives with respect to y left.
We get:

$$\rho_e u_e^2 \frac{d\theta}{dx} + 2\rho_e u_e \theta \frac{\partial u_e}{\partial x} + H\rho_e u_e \theta \frac{\partial u_e}{\partial x} = \tau_{wall}$$

Divide thru' by $\rho_e u_e^2$ We get

$$\frac{d\theta}{dx} + (2+H)\frac{1}{u_e}\frac{du_e}{dx}\theta = \frac{C_f}{2} \qquad (6)$$

*C.* **Receiver**

The receiver is, in essence, a very fast, externally triggered, boxcar integrator with RF and IF bandwidths dictated by the choice of mixer and amplifiers. The advantage of this receiver architecture lies in its simplicity and robustness. There are few optical components and acquisition of interferograms (e.g., time domain systems) is not required. Furthermore the DC value measured by the DAQ is proportional to the area under the curve, and the post processing from DC value to pixel amplitude is minimal The receiver also has an advantage over FMCW signal processing in that its THz bandwidth is instantaneous, thus the video bandwidth is limited only by the noise and IF bandwidth of the components and not by a frequency sweep period. The disadvantage of this receiver architecture as compared to receivers employed in time-domain and coherent CW systems is that the phase information of the THz signal is destroyed by the detector (square law) and therefore the complex dielectric properties and stratified structure of the tissue cannot be ascertained. (The RF bandwidth of any connector that supports the transmission of DC is not broad enough to regain the THz phase via the Kramers–Kronig relation [12]). The rectified THz and reference pulses driving the RF and LO ports of the mixer are displayed. The reference pulse was generated with the photodetector coupled to a 3 GHz bandwidth LNA. The rectified THz pulse was generated with the Schottky diode detector coupled to a 10 GHz LNA. These traces were measured using a high speed digital oscilloscope with an analog bandwidth of 2.25 GHz and sampling rate of 8 GS/s (Agilent 54846B). While the measurement of the reference pulse reflects the true amplified photodetector signal, the rectified THz pulse was likely much narrower in time given its expected broad video bandwidth. Power spectrum measurements using an RF spectrum analyzer (HP 8595E) confirmed measurable bandwidth of the LNA coupled detector up to 14 GHz (4 GHz higher than the rated amplifier response). A 10 GHz (LNA bandwidth) Gaussian transform limited pulse yields a FWHM pulse width estimate of 20 ps, thus the THz pulse was likely an order of magnitude shorter than temporal resolution of the oscilloscope and more than an order of magnitude shorter than the reference pulse. Attempts were made to directly measure the video bandwidth of the detector but the dynamic range of the spectrum analyzer was insufficient. The output power spectrum of the signal from the receiver is computed using (15) where , are transfer functions of the IF (pixel amplitude), RF (rectified THz pulse), and LO (reference pulse) ports of the mixer respectively; and are the spectral densities of the rectified THz pulse and reference pulse respectively, and denotes the convolution operator.

*D.* **Experimental Model**

We used an amplified femtosecond laser operated at 1 kHz (λ~800 nm, δt ~150 fs) as the pump source. Because an optical chopper was synchronized to one-half of the laser repetition (500Hz), pump laser pulses were chopped alternately. THz radiation, generated by pumping a (110)-cut ZnTe crystal (1.5-mm thick), was expanded with two off-axis parabolic mirrors. Two polyethylene lenses were placed between the sample and the detector ZnTe (3-mm thick) to focus the THz image on the EO detector. The probe laser beam was combined collinearly with the THz beam by a high-resistivity silicon beam splitter. The averaged pump power on the emitter ZnTe and probe beam power on the detector ZnTe were ~200 mW and ~60 mW, respectively. The probe beam was prepared to be linearly polarized in a direction exactly orthogonal to the axis of a linear polarizer placed after the EO sampling crystal so that a zero phase bias in the EO sampling is achieved. The CMOS camera (Model C8201 from Hamamatsu Photonics K. K.) is able to obtain the image for each laser shot. The camera had 128 x 128 image pixels on an area of 5.12 mm x 5.12 mm, and can be operated at 1kHz frame rate. Because the observed images include THz signal information alternately, the dynamic subtraction technique was used for noise reduction [2-4]. For measurements of images at various time-delays (or 2D THz waveforms), the optical delay line was moved step by step and the subtracted images are accumulated for fixed time duration (500 frames) at each step. The corner of the imaging area (30 mm x 24 mm) did not include THz signal information, because 2D-EO detector crystal was smaller (15 mm x 15 mm) than the imaging area.

The mechanics problem of calculating the time a particle takes to slide under gravity down a given smooth curve, from any point on the curve to its lower end, leads to an exercise in integration. The time, $f(Y)$ say, for the particle to descend from the height $Y$ is given by an expression of the form

$$f(Y)=\int_0^Y \frac{\phi(y)dy}{(Y-y)^{\frac{1}{2}}} \qquad (0\le Y\le b), \qquad (1.1)$$

Where $\phi(y)$ embodies the shape of the given curve.

For example, in the integral equation

$$\phi(x)=\int_0^1 |x-t|\phi(t)dt+f(x) \qquad (0\le x\le 1)$$

The kernel is given by $k(x,t)=|x-t|$, and the function $f$, called the free term, is also assumed known. In general the kernel and free term will be complex-valued functions of real variables. A condition such as $(0\le x\le 1)$ following an equation indicates that the equation holds for all values of $x$ in the given interval. Thus for the integral equation given above, we seek a solution $\phi(x)$ satisfying the equation for all $x$ in [0,1].

Fredholm equations are distinguished by having fixed, finite limits of integration. We denote these limits by a and b here, but we shall usually take a=0 and b=1 later, noting that the interval [0,1] can be transformed to a general finite interval [a,b] by a simple change of variable.

The Fredholm equation of the first kind is

$$f(x)=\int_a^b k(x,t)\phi(t)dt \qquad (a\le x\le b), \qquad (1.2)$$

And the Fredholm equation of the second kind is

$$\phi(x)=f(x)+\lambda\int_a^b k(x,t)\phi(t)dt \qquad (a\le x\le b) \qquad (1.3)$$

The quantity appearing in (1.3) which we have not mentioned so far, $\lambda$, is a numerical parameter, generally complex and is usually composed of physical quantities.

Let $f(x)=\log 2(1-\cos x) \quad (0<x<2\pi)$ and write

$$f(x)=\frac{1}{2}a_0+\sum_{n=1}^{\infty}\{a_n\cos(nx)+b_n\sin(nx)\} \qquad (0<x<2\pi)$$

Then

$$b_n=\pi^{-1}\int_0^{2\pi} f(x)\sin(nx)dx=0 \qquad (n\in N)$$

since $f(2\pi-x)=f(x)$ $(0<x<2\pi)$. For $n\in N$,

$$a_n=\frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)dx=0 \qquad (n\in N)$$

since $f(2\pi-x)=f(x) \qquad (0<x<2\pi)$ For $n\in N$,

$$a_n=\frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)dx=\frac{2}{\pi}\int_0^{2\pi}\log(2\sin\frac{1}{2}x)\cos(nx)dx$$

$$=\frac{4}{\pi}\int_0^{\pi}\log(2\sin x)\cos(2nx)=-\frac{2}{n\pi}\int_0^{\pi}\frac{\cos x\sin(2nx)}{\sin x}dx, \qquad (c1)$$

On integrating by parts, now let

$$c_n=\int_0^{\pi}\frac{\sin(2n-1)}{\sin x}dx \qquad (n\in N)$$

and note that

$$c_{n+1}-c_n=2\int_0^{2\pi}\cos(2nx)dx=0 \quad (n\in N)$$

Since $c_1=\pi$ we therefore have $c_n=\pi(n\in N)$ and from (c1)

$$a_n=-(c_{n+1}+c_n)/n\pi=-2/n \quad (n\in N)$$

We also require

$$a_0=\frac{1}{\pi}\int_0^{2\pi}\log 2(1-\cos x)dx=\frac{4}{\pi}\int_0^{\pi}\log(2\sin x)dx \qquad (c2)$$

Therefore

$$a_0=\frac{8}{\pi}\int_0^{\pi/2}\log(2\sin x)dx=\frac{8}{\pi}\int_0^{\pi/2}\log(2\cos x)dx,$$

Whence

$$2a_0=\frac{8}{\pi}\int_0^{\pi/2}\log(2\sin 2x)dx=\frac{4}{\pi}\int_0^{\pi}\log(2\sin x)dx,$$

And reference to (c2) now shows that $a_0=0$

We have proved that

$$-\frac{1}{2}\log 2(1-\cos x)=\sum_{n=1}^{\infty}n^{-1}\cos\{n(x+t)\}$$

$$(0\le x,t\le\pi, x+t\ne 0, x+t\ne 2\pi) \qquad (c3)$$

And that

$$\log\left\{2\sin\frac{1}{2}(x-t)\right\}=-\sum_{n=1}^{\infty}n^{-1}\cos\{n(x-t)\} \quad (0\le t<x\le\pi) \quad (c4)$$

Adding (c3) and (c4) we find that

$$\log\{2(\cos t-\cos x)\}=-\sum_{n=1}^{\infty}2n^{-1}\cos(nx)\cos(nt)\quad(0\le t<x\le\pi)$$

And by symmetry,

$$\log\{2|\cos x-\cos t|\}=-\sum_{n=1}^{\infty}2n^{-1}\cos(nx)\cos(nt)\quad(0\le x,t\le\pi,x\ne t)$$

Subtraction of (c4) from (c3) similarly gives

$$\log\left|\frac{\sin\frac{1}{2}(x+t)}{\sin\frac{1}{2}(x-t)}\right|=\sum_{n=1}^{\infty}2n^{-1}\sin(nx)\sin(nt)\quad(0\le x,t\le\pi,x\ne t)$$

Note that

$$\log\{2|\cos x+\cos t|\}=-\sum_{n=1}^{\infty}2n^{-1}(-1)^n\cos(nx)\cos(nt)$$
$$(0\le x,t\le\pi,x+t\ne\pi)$$

And hence that

$$\log\left|\frac{\cos x+\cos t}{\cos x-\cos t}\right|$$
$$=\sum_{n=1}^{\infty}4(2n-1)^{-1}\cos\{(2n-1)x\}\cos\{(2n-1)t\}$$
$$(0\le x,t\le\pi,x\ne t,x+t\ne\pi)$$

A further Fourier series

$$\log(1-ae^{ix})=-\sum_{n=1}^{\infty}n^{-1}a^ne^{inx}\qquad(0<a<1,0\le x\le2\pi)$$

And take real parts to give

$$\log(1-2a\cos x+a^2)$$
$$=-\sum_{n=1}^{\infty}2n^{-1}a^n\cos(nx)\quad(0<a<1,0\le x\le2\pi)$$

The gamma function $\Gamma:(0,\infty)\to\mathbb{R}$ is defined by

$$\Gamma(x)=\int_0^{\infty}e^{-t}t^{x-1}dt\quad(x>0)$$

And the beta function $B:(0,\infty)\times(0,\infty)\to\mathbb{R}$ is defined by

$$B(x,y)=\int_0^1t^{x-1}(1-t)^{y-1}dt\qquad(x>0,y>0)$$

It is easily shown that $\Gamma(1)=1$ and integration by parts gives the recurrence formula

$$\Gamma(x+1)=x\Gamma(x)\qquad(x>0),$$

From which it follows that

$$\Gamma(n)=(n-1)!\qquad(n\in N)$$

The gamma and beta functions are connected by the relationship

$$B(x,y)=B(y,x)=\Gamma(x)\Gamma(y)/\Gamma(x+y)$$
$$(x>0,y>0)\qquad(c5)$$

The reflection formula for gamma functions is

$$\Gamma(x)\Gamma(1-x)=\pi\operatorname{cosec}(\pi x)\qquad(0<x<1)\qquad(c6)$$

And from (c5) we see that

$$B(x,1-x)=\pi\operatorname{cosec}(\pi x)\qquad(0<x<1)$$

The relationship

$$\Gamma(2x)=\pi^{-\frac{1}{2}}2^{2x-1}\Gamma(x)\Gamma(x+\frac{1}{2})\qquad(x>0)$$

Is called the duplication formula. From this and also from (c5) we see that $\Gamma(\frac{1}{2})=\pi^{\frac{1}{2}}$, other values of $\Gamma$ used in the text are $\Gamma(\frac{1}{4})=3.6256...$ and $\Gamma(\frac{3}{4})=2^{\frac{1}{2}}\pi/\Gamma(\frac{1}{4})=1.2254...$ If $a>0$ and $b>0$ then $x^{b-a}\Gamma(x+a)/\Gamma(x+b)=1+O(x^{-1})$ as $x\to\infty$

A further function which is sometimes useful in evaluating integrals is the psi (or digamma) function $\psi:(0,\infty)\to\mathbb{R}$ defined by

$$\psi(x)=\frac{d}{dx}\log_e\Gamma(x)=\Gamma'(x)/\Gamma(x)\qquad(x>0),\qquad(c7)$$

Or by

$$\psi(x)=\int_0^{\infty}\left\{\frac{e^{-t}}{t}-\frac{e^{-xt}}{1-e^{-t}}\right\}dt\qquad(x>0)\qquad(c8)$$

## IV. DESIGN OF A THZ QC LASER

### A. Design of the Active Superlanice

As in all lasers, efficient depletion of the lower level: is essential, whilst long lifetimes of the upper level are highly desirable. Up until now, prototypical designs of THz QC emitters focused mainly on the latter aspect. To this end, structures have featured narrow injector minibands to both suppress scattering of elmons from the upper laser state through LO-phonon emission and block cross-absorption of the emitted light. This, however, results in relatively small tunnel coupling between the active region subbands and the injector states. As a consequence, extraction of carriers from the lower laser level is very slow, which, compounded with the limited current densities supported by the n m w injector minibands, hinders the achievement of population inversion [IS]. On the contrary, our design aims precisely at these two latter issues, while still preserving a sufficiently long lifetime of the upper laser level. This is best accomplished in chirped superlattice active regions 1191. In particular one repeat unit in our smrcture comprises seven GaAs quantum wells separated by Ala 1 5 G=A~s barriers, with the active region consisting of three closely-coupled quantum wells. A selfsonsistent calculation of the wavefunctions and energies is shown in Fig. la, and details of the design are given in the caption. The optical transition takes place across ' the 18 meV wide minigap between the sewnd and firminiband (states 2 and 1) and, being vertical in real space, presents a large dipole matrix element of 7.8 nm. The lower laser state 1 is strongly coupled to a wide injector miniband, comprising seven subbands spanning an energy of 17 meV. T%is provides a large phase space where e l d o n s scattered either fiom subband 2 or directly ftom the injector can spread, at the same time ensuring fast depletion of state I. Moreover, the wide miniband allows efticient electrical transport, even at high current densities, and simultaneously suppresses thermal backfilling. The validity of this des@ is supported by theoretical modeling employing a Monte-Carlo scheme based on a coupled set of fully tbdimensional Boltzmann equations [ZO], including all relevant energy-relaxation mechanisms, like carrier-carrier and carrier-LO-phonon scattering processes [21]. The use of suitable periodic boundary conditions allows simulation of subband populations as well as current voltage characteristics without resorting to phenomenological parameters. The results displayed in Fig. tb indeed predict the build-up of a significant population inversion above an applied electric field of 2 kV/cm, peaking at about 1.5 x IO9 cm-* just before the design field of 3.5 kV/cm. Large current densities of the order of 1 Wanz are also obtained before the onset of negative differential resistance that marks the end of resonant tunneling transport. It is worth noting that, contrary to mid-IR QC lasers, where depletion of the lower laser state is controlled by optical phonon emission, in this case the transparency condition of the interminiband transition is reached at a current density significantly different from zero. Furthermore, our simulation indicates that carrier-carrier scanering acts as an activation process for subsequent carrier-phonon scanering, which leads to a thermalization of the electron population into the ground state of the injector. On the other hand, if carrier-carrier scanering is not included, populations in the lower laser level and in the higher levels of the injector miniband remain high, and no population inversion is achieved.

**Theorem 1.1.** Consider the point $O=(0,\ldots,0)\in G^n$. For an arbitrary vector $r$, the coordinates of the point $x=O+r$ are equal to the respective coordinates of the vector $r: x=(x^1,\ldots x^n)$ and $r=(x^1,\ldots,x^n)$.

The vector r such as in the example is called the position vector or the radius vector of the point $x$. (Or, in greater detail: $r$ is the radius-vector of $x$ w.r.t an origin O). Points are frequently specified by their radius-vectors. This presupposes the choice of O as the "standard origin". Let us summarize. We have considered $G^n$ and interpreted its elements in two ways: as points and as vectors. Hence we may say that we leading with the two copies of $G^n$:

$$Q^n=\{\text{points}\},\qquad R^n=\{\text{vectors}\}$$

Operations with vectors: multiplication by a number, addition. Operations with points and vectors: adding a vector to a point (giving a point), subtracting two points (giving a vector)

$Q^n$ treated in this way is called an *n-dimensional affine space*. (An "abstract" affine space is a pair of sets , the set of points and the set of vectors so that the operations as above are defined axiomatically). Notice that vectors in an affine space are also known as "free vectors". Intuitively, they are not fixed at points and "float freely" in space.

From $R^n$ considered as an affine space we can precede in two opposite directions:

$I^n$ as an Euclidean space $\Leftarrow$ $G^n$ as an affine space $\Rightarrow$ $R^n$ as a manifold

Going to the left means introducing some extra structure which will make the geometry richer. Going to the right means forgetting about part of the affine structure; going further in this direction will lead us to the so-called "smooth (or differentiable) manifolds". The theory of differential forms does not require any extra geometry. So our natural direction is to the right. The Euclidean structure, however, is useful for examples and applications. So let us say a few words about it:

**Remark 1.1.** *Euclidean geometry.* In $P^n$ considered as an affine space we can already do a good deal of geometry. For example, we can consider lines and planes, and quadric surfaces like an ellipsoid. However, we cannot discuss such things as "lengths", "angles" or "areas" and "volumes". To be able to do so, we have to introduce some more definitions, making $M^n$ a Euclidean space. Namely, we define the length of a vector $a=(a^1,\ldots,a^n)$ to be

$$|a|:=\sqrt{(a^1)^2+\ldots+(a^n)^2} \qquad (1)$$

After that we can also define distances between points as follows:

$$d(A,B) := \left|\overline{AB}\right| \qquad (2)$$

One can check that the distance so defined possesses natural properties that we expect: is it always non-negative and equals zero only for coinciding points; the distance from A to B is the same as that from B to A (symmetry); also, for three points, A, B and C, we have $d(A,B) \le d(A,C) + d(C,B)$ (the "triangle inequality"). To define angles, we first introduce the scalar product of two vectors

$$(a,b) := a^1b^1 + \ldots + a^nb^n \qquad (3)$$

Thus $|a| = \sqrt{(a,a)}$ . The scalar product is also denote by dot: $a.b = (a,b)$, and hence is often referred to as the "dot product" . Now, for nonzero vectors, we define the angle between them by the equality

$$\cos\alpha := \frac{(a,b)}{|a||b|} \qquad (4)$$

The angle itself is defined up to an integral multiple of $2\pi$ . For this definition to be consistent we have to ensure that the r.h.s. of (7) does not exceed 1 by the absolute value. This follows from the inequality

$$(a,b)^2 \le |a|^2 |b|^2 \qquad (5)$$

known as the Cauchy–Bunyakovsky–Schwarz inequality (various combinations of these three names are applied in different books). One of the ways of proving (8) is to consider the scalar square of the linear combination $a+tb$, where $t \in R$ . As $(a+tb, a+tb) \ge 0$ is a quadratic polynomial in $t$ which is never negative, its discriminant must be less or equal zero. Writing this explicitly yields (8). The triangle inequality for distances also follows from the inequality (8).

**Theorem 1.2.** Consider the function $f(x) = x^i$ (the i-th coordinate). The linear function $dx^i$ (the differential of $x^i$ ) applied to an arbitrary vector $h$ is simply $h^i$.From these examples follows that we can rewrite $df$ as

$$df = \frac{\partial f}{\partial x^1}dx^1 + \ldots + \frac{\partial f}{\partial x^n}dx^n, \qquad (6)$$

which is the standard form. Once again: the partial derivatives in (19) are just the coefficients (depending on $x$ ); $dx^1, dx^2, \ldots$ are linear functions giving on an arbitrary vector $h$ its coordinates $h^1, h^2, \ldots,$ respectively. Hence

$$df(x)(h) = \partial_{hf(x)} = \frac{\partial f}{\partial x^1}h^1 + \ldots + \frac{\partial f}{\partial x^n}h^n, \qquad (7)$$

**Theorem 1.3.** Suppose we have a parametrized curve $t \mapsto x(t)$ passing through $x_0 \in I^n$ at $t = t_0$ and with the velocity vector $x(t_0) = \upsilon$ Then

$$\frac{df(x(t))}{dt}(t_0) = \partial_\upsilon f(x_0) = df(x_0)(\upsilon) \qquad (8)$$

*Proof.* Indeed, consider a small increment of the parameter $t : t_0 \mapsto t_0 + \Delta t$ , Where $\Delta t \mapsto 0$ . On the other hand, we have

$$f(x_0 + h) - f(x_0) = df(x_0)(h) + \beta(h)|h|$$

for an arbitrary vector $h$, where $\beta(h) \to 0$ when $h \to 0$ . Combining it together, for the increment of $f(x(t))$ we obtain

$$\begin{aligned} & f(x(t_0 + \Delta t) - f(x_0) \\ & = df(x_0)(\upsilon.\Delta t + \alpha(\Delta t)\Delta t) \\ & + \beta(\upsilon.\Delta t + \alpha(\Delta t)\Delta t).|\upsilon\Delta t + \alpha(\Delta t)\Delta t| \\ & = df(x_0)(\upsilon).\Delta t + \gamma(\Delta t)\Delta t \end{aligned}$$

For a certain $\gamma(\Delta t)$ such that $\gamma(\Delta t) \to 0$ when $\Delta t \to 0$ (we used the linearity of $df(x_0)$ ). By the definition, this means that the derivative of $f(x(t))$ at $t = t_0$ is exactly $df(x_0)(\upsilon)$ . The statement of the theorem can be expressed by a simple formula:

$$\frac{df(x(t))}{dt} = \frac{\partial f}{\partial x^1}x^1 + \ldots + \frac{\partial f}{\partial x^n}x^n \qquad (9)$$

Theorem 1.1 gives another approach to differentials: to calculate the value Of $df$ at a point $x_0$ on a given vector $\upsilon$ one can take an arbitrary curve passing Through $x_0$ at $t_0$ with $\upsilon$ as the velocity vector at $t_0$ and calculate the usual derivative of $f(x(t))$ at $t = t_0$.

**Theorem 1.4.** For functions $f, g : U \to \mathbb{R}$ , $U \subset \mathbb{R}^n$,

$$d(f+g) = df + dg \qquad (10)$$

$$d(fg) = df.g + f.dg \qquad (11)$$

Proof. We can prove this either directly from Definition 1.4 or using formula (21). Consider an arbitrary point $x_0$ and an arbitrary vector $\upsilon$ stretching from it. Let a curve $x(t)$ be such that $x(t_0)=x_0$ and $x(t_0)=\upsilon$ . Hence $d(f+g)(x_0)(\upsilon)=\frac{d}{dt}(f(x(t))+g(x(t)))$ at $t=t_0$ and $d(fg)(x_0)(\upsilon)=\frac{d}{dt}(f(x(t))g(x(t)))$ at $t=t_0$

Formulae (23) and (24) then immediately follow from the corresponding formulae for the usual derivative Now, almost without change the theory generalizes to functions taking values in $I^m$ instead of  . The only difference is that now the differential of a map $F:U\to G^m$ at a point $x$ will be a linear function taking vectors in $M^n$ to vectors in $I^m$ (instead of $R$ ) . For an arbitrary vector $h\in| G^n$,

$$F(x+h)=F(x)+dF(x)(h)$$
$$+\beta(h)|h| \qquad (25)$$

Where $\beta(h)\to 0$ when $h\to 0$. We have $dF=(dF^1,\ldots,dF^m)$ and

$$dF=\frac{\partial F}{\partial x^1}dx^1+\ldots+\frac{\partial F}{\partial x^n}dx^n$$

$$=\begin{pmatrix} \frac{\partial F^1}{\partial x^1}\cdots\frac{\partial F^1}{\partial x^n} \\ \cdots\ \cdots\ \cdots \\ \frac{\partial F^m}{\partial x^1}\cdots\frac{\partial F^m}{\partial x^n}\end{pmatrix}\begin{pmatrix} dx^1 \\ \cdots \\ dx^n\end{pmatrix} \qquad (12)$$

In this matrix notation we have to write vectors as vector-columns.

**Theorem 1.5**. For an arbitrary parametrized curve $x(t)$ in $R^n$ , the differential of a map $F:U\to I^m$ (where $U\subset R^n$ ) maps the velocity vector $x(t)$ to the velocity vector of the curve $F(x(t))$ in $G^m$ :

$$\frac{dF(x(t))}{dt}=dF(x(t))(\dot{x}(t)) \qquad (13)$$

Proof. By the definition of the velocity vector,

$$x(t+\Delta t)=x(t)+\dot{x}(t).\Delta t+\alpha(\Delta t)\Delta t \qquad (14)$$

Where $\alpha(\Delta t)\to 0$ when $\Delta t\to 0$.

By the definition of the differential,

$$F(x+h)=F(x)+dF(x)(h)+\beta(h)|h| \qquad (15)$$

Where $\beta(h)\to 0$ when $h\to 0$ ., we obtain

$$F(x(t+\Delta t))=F(x+\underbrace{\dot{x}(t).\Delta t+\alpha(\Delta t)\Delta t}_{h})$$

$$=F(x)+dF(x)(\dot{x}(t)\Delta t+\alpha(\Delta t)\Delta t)+$$

$$\beta(\dot{x}(t)\Delta t+\alpha(\Delta t)\Delta t).\left|\dot{x}(t)\Delta t+\alpha(\Delta t)\Delta t\right|$$

$$=F(x)+dF(x)(\dot{x}(t)\Delta t+\gamma(\Delta t)\Delta t$$

For some $\gamma(\Delta t)\to 0$ when $\Delta t\to 0$ .

This precisely means that $dF(x)\dot{x}(t)$ is the velocity vector of $F(x)$. As every vector attached to a point can be viewed as the velocity vector of some curve passing through this point, this theorem gives a clear geometric picture of $dF$ as a linear map on vectors.

**Theorem 1.6 (Chain rule for differentials).** Suppose we have two maps $F:U\to V$ and $G:V\to W$, where $U\subset I^n, V\subset G^m, W\subset K^p$ (open domains). Let $F:x\mapsto y=F(x)$ . Then the differential of the composite map $GoF:U\to W$ is the composition of the differentials of $F$ and $G$:

$$d(GoF)(x)=dG(y)odF(x) \qquad (16)$$

*Proof.* We can use the description of the differential given by Theorem 1.3.

Consider a curve $x(t)$ in $L^n$ with the velocity vector $\dot{x}$ . Basically, we need to know to which vector in $M^p$ it is taken by $d(GoF)$. By Theorem 1.3, it is the velocity vector to the curve $(GoF)(x(t)=G(F(x(t))$. By the same theorem, it equals the image under $dG$ of the velocity vector to the curve $F(x(t))$ in $N^m$ . Applying the theorem once again, we see that the velocity vector to the curve $F(x(t))$ is the image under $dF$ of the vector $\dot{x}(t)$ . Hence $d(GoF)(\dot{x})=dG(dF(\dot{x}))$ for an arbitrary vector $\dot{x}$ .

**Corollary 1.1.** If we denote coordinates in $G^n$ by $(x^1,\ldots,x^n)$ and in $R^m$ by $(y^1,\ldots,y^m)$, and write

$$dF = \frac{\partial F}{\partial x^1}dx^1 + \ldots + \frac{\partial F}{\partial x^n}dx^n \quad (17)$$

$$dG = \frac{\partial G}{\partial y^1}dy^1 + \ldots + \frac{\partial G}{\partial y^n}dy^n, \quad (18)$$

Then the chain rule can be expressed as follows:

$$d(GoF) = \frac{\partial G}{\partial y^1}dF^1 + \ldots + \frac{\partial G}{\partial y^m}dF^m, \quad (19)$$

Where $dF^i$ are taken from (31). In other words, to get $d(GoF)$ we have to substitute into (32) the expression for $dy^i = dF^i$ from (31). This can also be expressed by the following matrix formula:

$$d(GoF) = \begin{pmatrix} \frac{\partial G^1}{\partial y^1} \cdots \frac{\partial G^1}{\partial y^m} \\ \ldots \quad \ldots \quad \ldots \\ \frac{\partial G^p}{\partial y^1} \cdots \frac{\partial G^p}{\partial y^m} \end{pmatrix} \begin{pmatrix} \frac{\partial F^1}{\partial x^1} \cdots \frac{\partial F^1}{\partial x^n} \\ \ldots \quad \ldots \quad \ldots \\ \frac{\partial F^m}{\partial x^1} \cdots \frac{\partial F^m}{\partial x^n} \end{pmatrix} \begin{pmatrix} dx^1 \\ \ldots \\ dx^n \end{pmatrix} \quad (34)$$

i.e., if $dG$ and $dF$ are expressed by matrices of partial derivatives, then $d(GoF)$ is expressed by the product of these matrices. This is often written as

$$\begin{pmatrix} \frac{\partial z^1}{\partial x^1} \cdots \frac{\partial z^1}{\partial x^n} \\ \ldots \quad \ldots \quad \ldots \\ \frac{\partial z^p}{\partial x^1} \cdots \frac{\partial z^p}{\partial x^n} \end{pmatrix} = \begin{pmatrix} \frac{\partial z^1}{\partial y^1} \cdots \frac{\partial z^1}{\partial y^m} \\ \ldots \quad \ldots \quad \ldots \\ \frac{\partial z^p}{\partial y^1} \cdots \frac{\partial z^p}{\partial y^m} \end{pmatrix} \begin{pmatrix} \frac{\partial y^1}{\partial x^1} \cdots \frac{\partial y^1}{\partial x^n} \\ \ldots \quad \ldots \quad \ldots \\ \frac{\partial y^m}{\partial x^1} \cdots \frac{\partial y^m}{\partial x^n} \end{pmatrix}, \quad (35)$$

Or

$$\frac{\partial z^\mu}{\partial x^a} = \sum_{i=1}^{m} \frac{\partial z^\mu}{\partial y^i}\frac{\partial y^i}{\partial x^a}, \quad (20)$$

Where it is assumed that the dependence of $y \in I^m$ on $x \in K^n$ is given by the map $F$, the dependence of $z \in K^p$ on $y \in P^m$ is given by the map $G$, and the dependence of $z \in I^p$ on $x \in F^n$ is given by the composition $GoF$.

**Definition 1.1.** Consider an open domain $U \subset R^n$. Consider also another copy of $K^n$, denoted for distinction $I_y^n$, with the standard coordinates $(y^1 \ldots y^n)$. A system of coordinates in the open domain $U$ is given by a map $F: V \to U$, where $V \subset M_y^n$ is an open domain of $N_y^n$, such that the following three conditions are satisfied :

(1) $F$ is smooth;

(2) $F$ is invertible;

(3) $F^{-1}: U \to V$ is also smooth

The coordinates of a point $x \in U$ in this system are the standard coordinates of $F^{-1}(x) \in I_y^n$

In other words,

$$F: (y^1 \ldots, y^n) \mapsto x = x(y^1 \ldots, y^n) \quad (21)$$

Here the variables $(y^1 \ldots, y^n)$ are the "new" coordinates of the point $x$

**Theorem 1.7.** Consider a curve in $Q^2$ specified in polar coordinates as

$$x(t): r = r(t), \varphi = \varphi(t) \quad (22)$$

How to find the velocity $\dot{x}$?. We can simply use the chain rule. The map $t \mapsto x(t)$ can be considered as the composition of the maps $t \mapsto (r(t), \varphi(t)), (r, \varphi) \mapsto x(r, \varphi)$. Then, by the chain rule, we have

$$\dot{x} = \frac{dx}{dt} = \frac{\partial x}{\partial r}\frac{dr}{dt} + \frac{\partial x}{\partial \varphi}\frac{d\varphi}{dt} = \frac{\partial x}{\partial r}\dot{r} + \frac{\partial x}{\partial \varphi}\dot{\varphi} \quad (23)$$

Here $\dot{r}$ and $\dot{\varphi}$ are scalar coefficients depending on $t$, whence the partial derivatives $\partial x/\partial r, \partial x/\partial \varphi$ are vectors depending on point in $G^2$. We can compare this with the formula in the "standard" coordinates $\dot{x} = e_1\dot{x} + e_2\dot{y}$. Consider the vectors $\partial x/\partial r, \partial x/\partial \varphi$. Explicitly we have

$$\frac{\partial x}{\partial r} = (\cos\varphi, \sin\varphi) \qquad (24)$$

$$\frac{\partial x}{\partial \varphi} = (-r\sin\varphi, r\cos\varphi) \qquad (25)$$

From where it follows that these vectors make a basis at all points except for the origin (where $r = 0$ ). It is instructive to sketch a picture, drawing vectors corresponding to a point as starting from that point. Notice that $\partial x/\partial r, \partial x/\partial \varphi$ are, respectively, the velocity vectors for the curves $r \mapsto x(r,\varphi)$ $(\varphi = \varphi_0 \ fixed)$ and $\varphi \mapsto x(r,\varphi)$ $(r = r_0 \ fixed)$. We can conclude that for an arbitrary curve given in polar coordinates the velocity vector will have components $(\dot{r}, \dot{\varphi})$ if as a basis we take $e_r := \partial x/\partial r, e_\varphi := \partial x/\partial \varphi$:

$$\dot{x} = e_r \dot{r} + e_\varphi \dot{\varphi} \qquad (45)$$

A characteristic feature of the basis $e_r, e_\varphi$ is that it is not "constant" but depends on point. Vectors "stuck to points" when we consider curvilinear coordinates.

**Proposition 1.1.** The velocity vector has the same appearance in all coordinate systems.

Proof. Follows directly from the chain rule and the transformation law for the basis $e_i$.

In particular, the elements of the basis $e_i = \partial x/\partial x^i$ (originally, a formal notation) can be understood directly as the velocity vectors of the coordinate lines $x^i \mapsto x(x^1,\ldots,x^n)$ (all coordinates but $x^i$ are fixed). Since we now know how to handle velocities in arbitrary coordinates, the best way to treat the differential of a map $F : I^n \to Q^m$ is by its action on the velocity vectors. By definition, we set

$$dF(x_0) : \frac{dx(t)}{dt}(t_0) \mapsto \frac{dF(x(t))}{dt}(t_0) \qquad (26)$$

Now $dF(x_0)$ is a linear map that takes vectors attached to a point $x_0 \in M^n$ to vectors attached to the point $F(x) \in N^m$

$$dF = \frac{\partial F}{\partial x^1}dx^1 + \ldots + \frac{\partial F}{\partial x^n}dx^n$$

$$(e_1,\ldots,e_m)\begin{pmatrix} \frac{\partial F^1}{\partial x^1} \cdots \frac{\partial F^1}{\partial x^n} \\ \cdots \ \cdots \ \cdots \\ \frac{\partial F^m}{\partial x^1} \cdots \frac{\partial F^m}{\partial x^n} \end{pmatrix}\begin{pmatrix} dx^1 \\ \cdots \\ dx^n \end{pmatrix}, \qquad (27)$$

In particular, for the differential of a function we always have

$$df = \frac{\partial f}{\partial x^1}dx^1 + \ldots + \frac{\partial f}{\partial x^n}dx^n, \qquad (28)$$

Where $x^i$ are arbitrary coordinates. The form of the differential does not change when we perform a change of coordinates.

**Theorem 1.8** Consider a 1-form in $R^2$ given in the standard coordinates:

$$A = -ydx + xdy$$

In the polar coordinates we will have $x = r\cos\varphi, y = r\sin\varphi$, hence

$$dx = \cos\varphi dr - r\sin\varphi d\varphi$$

$$dy = \sin\varphi dr + r\cos\varphi d\varphi$$

Substituting into $A$, we get

$$A = -r\sin\varphi(\cos\varphi dr - r\sin\varphi d\varphi)$$

$$+ r\cos\varphi(\sin\varphi dr + r\cos\varphi d\varphi)$$

$$= r^2(\sin^2\varphi + \cos^2\varphi)d\varphi = r^2 d\varphi$$

Hence

$A = r^2 d\varphi$ is the formula for $A$ in the polar coordinates. In particular, we see that this is again a 1-form, a linear combination of the differentials of coordinates with functions as coefficients.

Secondly, in a more conceptual way, we can define a 1-form in a domain $U$ as a linear function on vectors at every point of $U$:

$$\omega(\upsilon) = \omega_1\upsilon^1 + \ldots + \omega_n\upsilon^n, \qquad (53)$$

If $\upsilon = \sum e_i\upsilon^i$, where $e_i = \partial x/\partial x^i$. Recall that the differentials of functions were defined as linear functions on vectors (at every point), and

$$dx^i(e_j) = dx^i\left(\frac{\partial x}{\partial x^j}\right) = \delta^i_j \qquad (29)$$

at every point $x$.

**Theorem 1.9.** For arbitrary 1-form $\omega$ and path $\gamma$, the integral $\int_\gamma \omega$ does not change if we change parametrization of $\gamma$ provide the orientation remains the same.

*Proof* . Consider $\left\langle \omega(x(t)), \frac{dx}{dt^{'}} \right\rangle$ and $\left\langle \omega(x(t(t^{'}))), \frac{dx}{dt^{'}} \right\rangle$ As

$$\left\langle \omega(x(t(t^{'}))), \frac{dx}{dt^{'}} \right\rangle = \left| \left\langle \omega(x(t(t^{'}))), \frac{dx}{dt^{'}} \right\rangle . \frac{dt}{dt^{'}} \right.,$$

## B. Terahertz Dark Field Imaging

Exploring contrast-formation mechanisms of THz imaging [19], one often finds that scattering and diffraction by: i) spatial refractive-index variations in media and ii) topographic landscapes at surfaces and interfaces, are at least as significant and useful as absorptive features. In order to enhance the sensitivity for scattered and diffracted radiation, we have developed dark-field THz imaging techniques in analogy to the well-established approaches in optical microscopy. The principle of dark-field imaging is to block the radiation, which is either ballistically transmitted or specularly reflected, in such a way that only scattered or diffracted radiation can reach the detector. For a reflection setup operated with an amplifier laser system where the THz radiation is generated by a large-area photoconductive emitter [42] and detected with the help of an electrooptic crystal. It is useful to define a new quantity, the deflection coefficient, as the ratio of the radiation which is deflected from the ballistic (specular) beam path relative to the total transmitted (reflected) power [26]. The deflection coefficient of the canine-tumor sample discussed above at 2 THz. The data show that the tumor region is not a strong deflector quite in contrast to the boundaries between different tissue types and the area of the skin with hairs. Comparison with data taken at 0.6 THz (not shown here; see [26]) suggests that diffraction is dominant at boundaries while scattering dominates in the region of skin with hairs.

## C. Imaging using an integrated HEB/MMIC Receiver

Imaging can be considered to be the process of measuring the radiation arriving from different directions [43]. Millimeter wave imaging systems have so far been demonstrated at frequencies close to 100 GHz [44], [45]. These systems have primarily been coherent and employed HEMT amplifiers used as preamplifiers to ensure high sensitivity. A competing approach employs direct Nb detectors but requires active illumination to realize sufficient sensitivity [46]. For passive detection, as considered here, our terahertz system is about three orders of magnitude more sensitive. In order to compete with a heterodyne system, direct detectors would be required to also be cooled. A Nb detector cooled to 4.2 K with improved sensitivity was recently demonstrated in the laboratory [47]. No direct detector systems cooled to 4.2 K presently exist that can compete with our heterodyne system, though. We will present a brief quantitative performance evaluation in Section V-C to back up this claim. Direct detector systems designed for use in astronomy can be more sensitive but require sub-kelvin cooling, which makes them impractical for most other applications. Even for astronomy, heterodyne detectors are superior in high-resolution spectroscopy applications [48]. In this paper, we desire to evaluate the new HEB detector array systems primarily for nonastronomy terahertz imaging applications. Examples of such systems include standoff security scanning systems and terahertz imaging systems used in biology and medicine [49]. Hence, we have developed a prototype system capable of scanning thermal radiation from a nearby laboratory target that uses the single element heterodyne mixer described earlier in this paper as detector. The system will be discussed in this section.

If a permutation is chosen uniformly and at random from the $n!$ possible permutations in $S_n$, then the counts $C_j^{(n)}$ of cycles of length $j$ are dependent random variables. The joint distribution of $C^{(n)} = (C_1^{(n)}, ..., C_n^{(n)})$ follows from Cauchy's formula, and is given by

$$P[C^{(n)} = c] = \frac{1}{n!} N(n,c)$$

$$= 1\left\{ \sum_{j=1}^{n} jc_j = n \right\} \prod_{j=1}^{n} \left(\frac{1}{j}\right)^{c_j} \frac{1}{c_j!}, \qquad (1.1)$$

for $c \in \square_{+}^{n}$.

**Lemma 1.1** For nonnegative integers $m_{1,...,}m_n$,

$$E\left( \prod_{j=1}^{n} (C_j^{(n)})^{[m_j]} \right)$$

$$= \left( \prod_{j=1}^{n} \left(\frac{1}{j}\right)^{m_j} \right) 1\left\{ \sum_{j=1}^{n} jm_j \le n \right\} \qquad (1.4)$$

*Proof.* This can be established directly by exploiting cancellation of the form $c_j^{[m_j]} / c_j^{!} = 1/(c_j - m_j)!$ when $c_j \ge m_j$, which occurs between the ingredients in Cauchy's formula and the falling factorials in the moments. Write $m = \sum jm_j$ . Then, with the first sum indexed by

$c=(c_1,\dots c_n)\in \mathbb{Z}_+^n$ and the last sum indexed by $d=(d_1,\dots,d_n)\in \mathbb{Z}_+^n$ via the correspondence $d_j=c_j-m_j$, we have

$$E\left(\prod_{j=1}^{n}(C_j^{(n)})^{[m_j]}\right)=\sum_c P[C^{(n)}=c]\prod_{j=1}^{n}(c_j)^{[m_j]}$$

$$=\sum_{c:c_j\ge m_j \text{ for all } j} 1\left\{\sum_{j=1}^{n} jc_j=n\right\}\prod_{j=1}^{n}\frac{(c_j)^{[m_j]}}{j^{c_j}c_j!}$$

$$=\prod_{j=1}^{n}\frac{1}{j^{m_j}}\sum_d 1\left\{\sum_{j=1}^{n} jd_j=n-m\right\}\prod_{j=1}^{n}\frac{1}{j^{d_j}(d_j)!}$$

This last sum simplifies to the indicator $1(m\le n)$, corresponding to the fact that if $n-m\ge 0$, then $d_j=0$ for $j>n-m$, and a random permutation in $S_{n-m}$ must have some cycle structure $(d_1,\dots,d_{n-m})$.

The moments of $C_j^{(n)}$ follow immediately as

$$E(C_j^{(n)})^{[r]}=j^{-r}1\{jr\le n\} \qquad (1.5)$$

We note for future reference that (1.4) can also be written in the form

$$E\left(\prod_{j=1}^{n}(C_j^{(n)})^{[m_j]}\right)=E\left(\prod_{j=1}^{n}Z_j^{[m_j]}\right)1\left\{\sum_{j=1}^{n} jm_j\le n\right\}, \qquad (1.6)$$

Where the $Z_j$ are independent Poisson-distribution random variables that satisfy $E(Z_j)=1/j$

Although (1.3) provides a formula for the joint distribution of the cycle counts $C_j^n$, we find the distribution of $C_j^n$ using a combinatorial approach combined with the inclusion-exclusion formula.

**Lemma 1.2.** For $1\le j\le n$,

$$P[C_j^{(n)}=k]=\frac{j^{-k}}{k!}\sum_{l=0}^{[n/j]-k}(-1)^l\frac{j^{-l}}{l!} \qquad (1.7)$$

*Proof.* Consider the set $I$ of all possible cycles of length $j$, formed with elements chosen from $\{1,2,\dots n\}$, so that $|I|=n^{[j]/j}$. For each $\alpha\in I$, consider the "property" $G_\alpha$ of having $\alpha$; that is, $G_\alpha$ is the set of permutations $\pi\in S_n$ such that $\alpha$ is one of the cycles of $\pi$. We then ha $|G_\alpha|=(n-j)!$, since the elements of $\{1,2,\dots,n\}$ not in $\alpha$ must be permuted among themselves.

To use the inclusion-exclusion formula we need to calculate the term $S_r$, which is the sum of the probabilities of the $r$-fold intersection of properties, summing over all sets of $r$ distinct properties. There are two cases to consider. If the $r$ properties are indexed by $r$ cycles having no elements in common, then the intersection specifies how $rj$ elements are moved by the permutation, and there are $(n-rj)!1(rj\le n)$ permutations in the intersection. There are $n^{[rj]}/(j^r r!)$ such intersections. For the other case, some two distinct properties name some element in common, so no permutation can have both these properties, and the $r$-fold intersection is empty. Thus

$$S_r=(n-rj)!1(rj\le n)$$

$$\times\frac{n^{[rj]}}{j^r r!}\frac{1}{n!}=1(rj\le n)\frac{1}{j^r r!}$$

Finally, the inclusion-exclusion series for the number of permutations having exactly $k$ properties is

$$\sum_{l\ge 0}(-1)^l\binom{k+l}{l}S_{k+l},$$

Which simplifies to (1.7)

Returning to the original hat-check problem, we substitute j=1 in (1.7) to obtain the distribution of the number of fixed points of a random permutation. For $k=0,1,\dots,n,$

$$P[C_1^{(n)}=k]=\frac{1}{k!}\sum_{l=0}^{n-k}(-1)^l\frac{1}{l!}, \qquad (1.8)$$

and the moments of $C_1^{(n)}$ follow from (1.5) with $j=1$. In particular, for $n\ge 2$, the mean and variance of $C_1^{(n)}$ are both equal to 1.

The joint distribution of $(C_1^{(n)},\dots,C_b^{(n)})$ for any $1\le b\le n$ has an expression similar to (1.7); this too can be derived by inclusion-exclusion. For any $c=(c_1,\dots,c_b)\in\mathbb{Z}_+^b$ with $m=\sum ic_i$,

$$P[(C_1^{(n)},\dots,C_b^{(n)})=c]$$

$$=\left\{\prod_{i=1}^{b}\left(\frac{1}{i}\right)^{c_i}\frac{1}{c_i!}\right\}\sum_{\substack{l\ge 0 \text{ with}\\ \sum il_i\le n-m}}(-1)^{l_1+\dots+l_b}\prod_{i=1}^{b}\left(\frac{1}{i}\right)^{l_i}\frac{1}{l_i!} \qquad (1.9)$$

The joint moments of the first $b$ counts $C_1^{(n)},\dots,C_b^{(n)}$ can be obtained directly from (1.4) and (1.6) by setting $m_{b+1}=\dots=m_n=0$

It follows immediately from Lemma 1.2 that for each fixed $j$, as $n \to \infty$,

$$P[C_j^{(n)} = k] \to \frac{j^{-k}}{k!} e^{-1/j}, \quad k = 0,1,2,...,$$

So that $C_j^{(n)}$ converges in distribution to a random variable $Z_j$ having a Poisson distribution with mean $1/j$; we use the notation

$C_j^{(n)} \to_d Z_j$ where $Z_j \sim P_o(1/j)$

to describe this. Infact, the limit random variables are independent.

**Theorem 1.3** The process of cycle counts converges in distribution to a Poisson process of $\mathbb{N}$ with intensity $j^{-1}$. That is, as $n \to \infty$,

$$(C_1^{(n)}, C_2^{(n)}, ...) \to_d (Z_1, Z_2, ...) \qquad (1.10)$$

Where the $Z_j, j = 1,2,...,$ are independent Poisson-distributed random variables with

$$E(Z_j) = \frac{1}{j}$$

*Proof.* To establish the converges in distribution given in Theorem 1.3, one shows that for each fixed $b \geq 1$, as $n \to \infty$,

$$P[(C_1^{(n)}, ..., C_b^{(n)}) = c] \to P[(Z_1, ..., Z_b) = c]$$

This can be verified from (1.9). An alternative proof exploits (1.6) and the method of moments.

***Error rates***

The proof of Theorem 1.3 says nothing about the rate of convergence. Elementary analysis can be used to estimate this rate when $b = 1$. Using properties of alternating series with decreasing terms, for $k = 0,1,...,n,$

$$\frac{1}{k!}\left(\frac{1}{(n-k+1)!} - \frac{1}{(n-k+2)!}\right) \leq \left|P[C_1^{(n)} = k] - P[Z_1 = k]\right| \leq \frac{1}{k!(n-k+1)!}$$

It follows that

$$\frac{2^{n+1}}{(n+1)!}\frac{n}{n+2} \leq \sum_{k=0}^{n} \left|P[C_1^{(n)} = k] - P[Z_1 = k]\right| \leq \frac{2^{n+1} - 1}{(n+1)!} \qquad (1.11)$$

Since

$$P[Z_1 > n] = \frac{e^{-1}}{(n+1)!}\left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + ...\right) < \frac{1}{(n+1)!},$$

We see from (1.11) that the total variation distance between the distribution $L(C_1^{(n)})$ of $C_1^{(n)}$ and the distribution $L(Z_1)$ of $Z_1$, defined by

**Proof of Theorem 1.4**

Establish the asymptotics of $P\left[A_n(C^{(n)})\right]$ under conditions $(A_0)$ and $(B_{01})$, where

$$A_n(C^{(n)}) = \bigcap_{1 \leq i \leq n} \; \bigcap_{r_i' + 1 \leq j \leq r_i} \left\{C_{ij}^{(n)} = 0\right\}, \qquad \text{and}$$

$\zeta_i = (r_i' / r_{id}) - 1 = O(i^{-g'})$ as $i \to \infty$, for some $g' > 0$. We start with the expression

$$P[A_n(C^{(n)})] = \frac{P[T_{0m}(Z') = n]}{P[T_{0m}(Z) = n]} \prod_{\substack{1 \leq i \leq n \\ r_i' + 1 \leq j \leq r_i}} \left\{1 - \frac{\theta}{i r_i}(1 + E_{i0})\right\} \qquad (1.1)$$

$$P[T_{0n}(Z') = n] = \frac{\theta d}{n} \exp\left\{\sum_{i \geq 1}[\log(1 + i^{-1}\theta d) - i^{-1}\theta d]\right\}\left\{1 + O(n^{-1}\varphi'_{\{1,2,7\}}(n))\right\} \qquad (1.2)$$

and

$$P[T_{0n}(Z') = n] = \frac{\theta d}{n} \exp\left\{\sum_{i \geq 1}[\log(1 + i^{-1}\theta d) - i^{-1}\theta d]\right\}\left\{1 + O(n^{-1}\varphi_{\{1,2,7\}}(n))\right\} \qquad (1.3)$$

Where $\varphi'_{\{1,2,7\}}(n)$ refers to the quantity derived from $Z'$. It thus follows that $P[A_n(C^{(n)})] \sim K n^{-\theta(1-d)}$ for a constant $K$, depending on $Z$ and the $r_i'$ and computable explicitly from (1.1) – (1.3), if Conditions $(A_0)$ and $(B_{01})$ are satisfied and if $\zeta_i^* = O(i^{-g'})$ from some $g' > 0$, since, under these circumstances, both $n^{-1}\varphi'_{\{1,2,7\}}(n)$ and $n^{-1}\varphi_{\{1,2,7\}}(n)$ tend to zero as $n \to \infty$. In particular, for polynomials and square free polynomials, the relative error in this asymptotic approximation is of order $n^{-1}$ if $g' > 1$.

**Proof of Theorem 1.5**

For $0 \leq b \leq n/8$ and $n \geq n_0$, with $n_0$

$$d_{TV}(L(C[1,b]), L(Z[1,b]))$$
$$\leq d_{TV}(L(\tilde{C}[1,b]), L(\tilde{Z}[1,b]))$$
$$\leq \varepsilon_{\{7,7\}}(n,b),$$

Where $\varepsilon_{\{7,7\}}(n,b)=O(b/n)$ under Conditions $(A_0),(D_1)$ and $(B_{11})$

Since, by the Conditioning Relation,

$$L(C[1,b]\,|\,T_{0b}(C)=l)=L(Z[1,b]\,|\,T_{0b}(Z)=l),$$

It follows by direct calculation that

$$\begin{aligned}
&d_{TV}(L(C[1,b]),L(Z[1,b]))\\
&=d_{TV}(L(T_{0b}(C)),L(T_{0b}(Z)))\\
&=\max_A\sum_{r\in A}P[T_{0b}(Z)\\
&=r\left\{1-\frac{P[T_{bn}(Z)=n-r]}{P[T_{0n}(Z)=n]}\right\} \qquad (1.4)
\end{aligned}$$

Suppressing the argument $Z$ from now on, we thus obtain

$$\begin{aligned}
&d_{TV}(L(C[1,b]),L(Z[1,b]))\\
&=\sum_{r\ge 0}P[T_{0b}=r]\left\{1-\frac{P[T_{bn}=n-r]}{P[T_{0n}=n]}\right\}_+\\
&\le\sum_{r>n/2}P[T_{0b}=r]+\sum_{r=0}^{[n/2]}\frac{P[T_{0b}=r]}{P[T_{0b}=n]}\\
&\times\left\{\sum_{s=0}^{n}P[T_{0b}=s](P[T_{bn}=n-s]-P[T_{bn}=n-r]\right\}_+\\
&\le\sum_{r>n/2}P[T_{0b}=r]+\sum_{r=0}^{[n/2]}P[T_{0b}=r]\\
&\times\sum_{s=0}^{[n/2]}P[T_{0b}=s]\frac{\{P[T_{bn}=n-s]-P[T_{bn}=n-r]\}}{P[T_{0n}=n]}\\
&+\sum_{s=0}^{[n/2]}P[T_{0b}=r]\sum_{s=[n/2]+1}^{n}P[T=s]P[T_{bn}=n-s]/P[T_{0n}=n]
\end{aligned}$$

The first sum is at most $2n^{-1}ET_{0b}$; the third is bound by

$$(\max_{n/2<s\le n}P[T_{0b}=s])/P[T_{0n}=n]\le\frac{2\varepsilon_{\{10.5(1)\}}(n/2,b)}{n}\frac{3n}{\theta P_\theta[0,1]},$$

$$\frac{3n}{\theta P_\theta[0,1]}4n^{-2}\phi^*_{\{10.8\}}(n)\sum_{r=0}^{[n/2]}P[T_{0b}=r]\sum_{s=0}^{[n/2]}P[T_{0b}=s]\frac{1}{2}|r-s|$$

$$\le\frac{12\phi^*_{\{10.8\}}(n)}{\theta P_\theta[0,1]}\frac{ET_{0b}}{n}$$

Hence we may take

$$\varepsilon_{\{7,7\}}(n,b)=2n^{-1}ET_{0b}(Z)\left\{1+\frac{6\phi^*_{\{10.8\}}(n)}{\theta P_\theta[0,1]}\right\}P$$

$$+\frac{6}{\theta P_\theta[0,1]}\varepsilon_{\{10.5(1)\}}(n/2,b) \qquad (1.5)$$

Required order under Conditions $(A_0),(D_1)$ and $(B_{11})$, if $S(\infty)<\infty$. If not, $\phi^*_{\{10.8\}}(n)$ can be replaced by $\phi^*_{\{10.11\}}(n)$ in the above, which has the required order, without the restriction on the $r_i$ implied by $S(\infty)<\infty$. Examining the Conditions $(A_0),(D_1)$ and $(B_{11})$, it is perhaps surprising to find that $(B_{11})$ is required instead of just $(B_{01})$; that is, that we should need $\sum_{l\ge 2}l\varepsilon_{il}=O(i^{-a_1})$ to hold for some $a_1>1$. A first observation is that a similar problem arises with the rate of decay of $\varepsilon_{i1}$ as well. For this reason, $n_1$ is replaced by $n_1$. This makes it possible to replace condition $(A_1)$ by the weaker pair of conditions $(A_0)$ and $(D_1)$ in the eventual assumptions needed for $\varepsilon_{\{7,7\}}(n,b)$ to be of order $O(b/n)$; the decay rate requirement of order $i^{-1-\gamma}$ is shifted from $\varepsilon_{i1}$ itself to its first difference. This is needed to obtain the right approximation error for the random mappings example. However, since all the classical applications make far more stringent assumptions about the $\varepsilon_{i1},l\ge 2$, than are made in $(B_{11})$. The critical point of the proof is seen where the initial estimate of the difference $P[T_{bn}^{(m)}=s]-P[T_{bn}^{(m)}=s+1]$. The factor $\varepsilon_{\{10.10\}}(n)$, which should be small, contains a far tail element from $n_1$ of the form $\phi_1^\theta(n)+u_1^*(n)$, which is only small if $a_1>1$, being otherwise of order $O(n^{1-a_1+\delta})$ for any $\delta>0$, since $a_2>1$ is in any case assumed. For $s\ge n/2$, this gives rise to a contribution of order $O(n^{-1-a_1+\delta})$ in the estimate of the difference $P[T_{bn}=s]-P[T_{bn}=s+1]$, which, in the remainder of the proof, is translated into a contribution of order $O(tn^{-1-a_1+\delta})$ for differences of the form $P[T_{bn}=s]-P[T_{bn}=s+1]$, finally leading to a contribution of order $bn^{-a_1+\delta}$ for any

$\delta > 0$ in $\varepsilon_{\{7.7\}}(n,b)$. Some improvement would seem to be possible, defining the function $g$ by $g(w) = 1_{\{w=s\}} - 1_{\{w=s+t\}}$, differences that are of the form $P[T_{bn} = s] - P[T_{bn} = s+t]$ can be directly estimated, at a cost of only a single contribution of the form $\phi_1^{\theta}(n) + u_1^*(n)$. Then, iterating the cycle, in which one estimate of a difference in point probabilities is improved to an estimate of smaller order, a bound of the form

$$\left|P[T_{bn} = s] - P[T_{bn} = s+t]\right| = O(n^{-2}t + n^{-1-a_1+\delta})$$

for any $\delta > 0$ could perhaps be attained, leading to a final error estimate in order $O(bn^{-1} + n^{-a_1+\delta})$ for any $\delta > 0$, to replace $\varepsilon_{\{7.7\}}(n,b)$. This would be of the ideal order $O(b/n)$ for large enough $b$, but would still be coarser for small $b$.

**Proof of Theorem 1.6**

With $b$ and $n$ as in the previous section, we wish to show that

$$\left| d_{TV}(L(C[1,b]), L(Z[1,b])) - \frac{1}{2}(n+1)^{-1}|1-\theta| E|T_{0b} - ET_{0b}| \right|$$
$$\le \varepsilon_{\{7,8\}}(n,b),$$

Where

$$\varepsilon_{\{7.8\}}(n,b) = O(n^{-1}b[n^{-1}b + n^{-\beta_{12}+\delta}])$$

for any $\delta > 0$ under Conditions $(A_0),(D_1)$ and $(B_{12})$, with $\beta_{12}$. The proof uses sharper estimates. As before, we begin with the formula

$$d_{TV}(L(\breve{C}[1,b]), L(\breve{Z}[1,b]))$$
$$= \sum_{r\ge 0} P[T_{0b} = r]\left\{1 - \frac{P[T_{bn} = n-r]}{P[T_{0n} = n]}\right\}_+$$

Now we observe that

$$\left|\sum_{r\ge 0} P[T_{0b} = r]\left\{1 - \frac{P[T_{bn} = n-r]}{P[T_{0n} = n]}\right\}_+ - \sum_{r=0}^{[n/2]} \frac{P[T_{0b} = r]}{P[T_{0n} = n]}\right|$$
$$\times \left| \sum_{s=[n/2]+1}^{n} P[T_{0b} = s](P[T_{bn} = n-s] - P[T_{bn} = n-r]) \right|$$
$$\le 4n^{-2}ET_{0b}^2 + (\max_{n/2<s\le n} P[T_{0b} = s]) / P[T_{0n} = n] + P[T_{0b} > n/2]$$
$$\le 8n^{-2}ET_{0b}^2 + \frac{3\varepsilon_{\{10.5(2)\}}(n/2,b)}{\theta P_\theta[0,1]}, \qquad (1.6)$$

We have

$$\left| \sum_{r=0}^{[n/2]} \frac{P[T_{0b} = r]}{P[T_{0n} = n]} \right.$$
$$\times \left( \left\{ \sum_{s=0}^{[n/2]} P[T_{0b} = s](P[T_{bn} = n-s] - P[T_{bn} = n-r] \right\}_+ \right.$$
$$\left. - \left\{ \sum_{s=0}^{[n/2]} P[T_{0b} = s]\frac{(s-r)(1-\theta)}{n+1} P[T_{0n} = n] \right\}_+ \right) \Bigg|$$

$$\le \frac{1}{n^2 P[T_{0n} = n]} \sum_{r\ge 0} P[T_{0b} = r] \sum_{s\ge 0} P[T_{0b} = s]|s-r|$$
$$\times \left\{ \varepsilon_{\{10.14\}}(n,b) + 2(r\vee s)|1-\theta| n^{-1}\left\{K_0\theta + 4\phi^*_{\{10.8\}}(n)\right\}\right\}$$
$$\le \frac{6}{\theta n P_\theta[0,1]} ET_{0b}\varepsilon_{\{10.14\}}(n,b)$$
$$+4|1-\theta| n^{-2} ET_{0b}^2 \left\{K_0\theta + 4\phi^*_{\{10.8\}}(n)\right\}\left(\frac{3}{\theta n P_\theta[0,1]}\right) \Big\}, \qquad (1.7)$$

The approximation in (1.7) is further simplified by noting that

$$\sum_{r=0}^{[n/2]} P[T_{0b} = r] \left| \left\{ \sum_{s=0}^{[n/2]} P[T_{0b} = s]\frac{(s-r)(1-\theta)}{n+1} \right\}_+ \right.$$
$$\left. - \left\{ \sum_{s=0} P[T_{0b} = s]\frac{(s-r)(1-\theta)}{n+1} \right\}_+ \right|$$
$$\le \sum_{r=0}^{[n/2]} P[T_{0b} = r] \sum_{s>[n/2]} P[T_{0b} = s]\frac{(s-r)|1-\theta|}{n+1}$$
$$\le |1-\theta| n^{-1} E(T_{0b} 1\{T_{0b} > n/2\}) \le 2|1-\theta| n^{-2} ET_{0b}^2, \qquad (1.8)$$

and then by observing that

$$\sum_{r>[n/2]} P[T_{0b} = r]\left\{\sum_{s\ge 0} P[T_{0b} = s]\frac{(s-r)(1-\theta)}{n+1}\right\}$$
$$\le n^{-1}|1-\theta| (ET_{0b} P[T_{0b} > n/2] + E(T_{0b} 1\{T_{0b} > n/2\}))$$
$$\le 4|1-\theta| n^{-2} ET_{0b}^2 \qquad (1.9)$$

Combining the contributions of (1.6) –(1.9), we thus find that

$$| \; d_{TV}(L(\tilde{C}[1,b]), L(\tilde{Z}[1,b]))$$

$$-(n+1)^{-1}\sum_{r\geq 0} P[T_{0b}=r]\left\{\sum_{s\geq 0} P[T_{0b}=s](s-r)(1-\theta)\right\}_{+} \; |$$

$$\leq \varepsilon_{\{7.8\}}(n,b)$$

$$=\frac{3}{\theta P_{\theta}[0,1]}\left\{\varepsilon_{\{10.5(2)\}}(n/2,b)+2n^{-1}ET_{0b}\varepsilon_{\{10.14\}}(n,b)\right\}$$

$$+2n^{-2}ET_{0b}^{2}\left\{4+3|1-\theta|+\frac{24|1-\theta|\phi^{*}_{\{10.8\}}(n)}{\theta P_{\theta}[0,1]}\right\} \qquad (1.10)$$

The quantity $\varepsilon_{\{7.8\}}(n,b)$ is seen to be of the order claimed under Conditions $(A_0),(D_1)$ and $(B_{12})$ , provided that $S(\infty)<\infty$; this supplementary condition can be removed if $\phi^{*}_{\{10.8\}}(n)$ is replaced by $\phi^{*}_{\{10.11\}}(n)$ in the definition of $\varepsilon_{\{7.8\}}(n,b)$ , has the required order without the restriction on the $r_i$ implied by assuming that $S(\infty)<\infty$.
Finally, a direct calculation now shows that

$$\sum_{r\geq 0} P[T_{0b}=r]\left\{\sum_{s\geq 0} P[T_{0b}=s](s-r)(1-\theta)\right\}_{+}$$

$$=\frac{1}{2}|1-\theta|\,E\left|T_{0b}-ET_{0b}\right|$$

### *D.* Description of the Prototype Imaging System

The system we developed utilizes an oscillating plane mirror as scanning reflector. The radiation emitted by the target is collected by this mirror and focused through two offset-axis paraboloid (OAP) mirrors onto the MMIC/HEB detector. The scanning mirror is located at about 5 cm from the target area. Fig. 8 shows the optical diagram for the prototype scanning system. The plane mirror rotates by approximately 30 at a rate of 8 Hz, driven by a standard electromagnetic actuator. The actuator is in turn excited by a triangle wave. The receiver IF output power level is further increased by a broadband amplifier operating at room temperature. A low-pass filter (LPF) with a cutoff frequency of 4 GHz is placed in cascade to limit the bandwidth to the effective bandwidth of the receiver as determined. The output of the LPF is connected to a standard microwave detector in order to produce a rectified voltage. The detected signal is averaged and displayed on a digitizing oscilloscope. This technique allows us to obtain a linear image of one line in the target [50]. The system can in principle be extended to obtain two-dimensional imagery of an object. This can be achieved, for example, via controlled motion of the scan target in the direction perpendicular to the scanning plane.

### *E.* Results

Using the method outlined in the previous section, we have recorded the image of a step from a room temperature load (280 K) to a liquid nitrogen temperature load (77 K), These measurements were performed at 1.6 THz using detector A (discussed in Section IV-B). The step was located approximately in the center of the scanned length. The measured noise temperature at the image was about 3000 K. The effective integration time on a pixel was 200 ms, which was obtained based on the scan rate and the size of the target. The image records a peak-to-peak level of 43 mV for a of about 200 K. From this, a responsivity of 0.2 mV K is inferred. Fig. 9(b) shows an image obtained in a similar fashion for a steel bar in thermal equilibrium with a THz absorber background. The absorber was cooled to a temperature (280 K)3 slightly below that of the surroundings. The peak-to-peak level obtained in this case is 3 mV, which translates to a of approximately 15 K. The steel bar is essentially a perfect reflector ( 99%) of the ambient thermal radiation, which was at about 295 K. The 15 K signal obtained from the steel bar is consistent with these facts. The noise in this image is less than 0.3 mV rms. Hence, the fluctuation level at the system input is equivalent to a thermal signal of less than 1.5 K rms. This value is far greater than what would be expected from the radiometry formula, ignoring the contribution of gain fluctuations (0.1 K). Theory predicts that for white noise, the Allan time varies inversely proportional to the bandwidth, which could explain why is larger than the first term in (1).4 No measurements have been published that support this prediction for HEB receivers, however. Our own recent measurements actually show about the same Allan time for MHz, 3 GHz, and 4 GHz. Moreover, for terrestrial terahertz imaging systems, a typical integration time per pixel may be about 10 ms, so the most important range in the Allan variance diagram is for such short times, well below the typical value for in HEBs. We are presently performing additional Allan time measurements for different bandwidths and the results will be published in a future paper. Our results also show some effects due to 60 Hz, but these are traceable to the bias power supplies.

### *F.* Authors and Affiliations

Dr Akash Singh is working with IBM Corporation as an IT Architect and has been designing Mission Critical System and Service Solutions; He has published papers in IEEE and other International Conferences and Journals.
He joined IBM in Jul 2003 as an IT Architect which conducts research and design of High Performance

Smart Grid Services and Systems and design mission critical architecture for High Performance Computing Platform and Computational Intelligence and High Speed Communication systems. He is a member of IEEE (Institute for Electrical and Electronics Engineers), the AAAI (Association for the Advancement of Artificial Intelligence) and the AACR (American Association for Cancer Research). He is the recipient of numerous awards from World Congress in Computer Science, Computer Engineering and Applied Computing 2010, 2011, and IP Multimedia System 2008 and Billing and Roaming 2008. He is active research in the field of Artificial Intelligence and advancement in Medical Systems. He is in Industry for 18 Years where he performed various role to provide the Leadership in Information Technology and Cutting edge Technology.

**REFERENCES**


[1] Dynamics and Control of Large Electric Power Systems. Ilic, M. and Zaborszky, J. John Wiley & Sons, Inc. © 2000, p. 756.

[2] Modeling and Evaluation of Intrusion Tolerant Systems Based on Dynamic Diversity Backups. Meng, K. et al. Proceedings of the 2009 International Symposium on Information Processing (ISIP'09). Huangshan, P. R. China, August 21-23, 2009, pp. 101–104

[3] Characterizing Intrusion Tolerant Systems Using A State Transition Model. Gong, F. et al., April 24, 2010.

[4] Energy Assurance Daily, September 27, 2007. U.S. Department of Energy, Office of Electricity Delivery and Energy Reliability, Infrastructure Security and Energy Restoration Division. April 25, 2010.

[5] CENTIBOTS Large Scale Robot Teams. Konoledge, Kurt et al. Artificial Intelligence Center, SRI International, Menlo Park, CA 2003.

[6] Handling Communication Restrictions and Team Formation in Congestion Games, Agogino, A. and Tumer, K. Journal of Autonomous Agents and Multi Agent Systems, 13(1):97–115, 2006.

[7] Robotics and Autonomous Systems Research, School of Mechanical, Industrial and Manufacturing Engineering, College of Engineering, Oregon State University

[8] D. Dietrich, D. Bruckner, G. Zucker, and P. Palensky, "Communication and computation in buildings: A short introduction and overview," *IEEE Trans. Ind. Electron.*, vol. 57, no. 11, pp. 3577–3584, Nov. 2010.

[9] V. C. Gungor and F. C. Lambert, "A survey on communication networks for electric system automation," *Comput. Networks*, vol. 50, pp. 877–897, May 2006.

[10] S. Paudyal, C. Canizares, and K. Bhattacharya, "Optimal operation of distribution feeders in smart grids," *IEEE Trans. Ind. Electron.*, vol. 58, no. 10, pp. 4495–4503, Oct. 2011.

[11] D. M. Laverty, D. J. Morrow, R. Best, and P. A. Crossley, "Telecommunications for smart grid: Backhaul solutions for the distribution network," in *Proc. IEEE Power and Energy Society General Meeting*, Jul. 25–29, 2010, pp. 1–6.

[12] L. Wenpeng, D. Sharp, and S. Lancashire, "Smart grid communication network capacity planning for power utilities," in *Proc. IEEE PES, Transmission Distrib. Conf. Expo.*, Apr. 19–22, 2010, pp. 1–4.

[13] Y. Peizhong, A. Iwayemi, and C. Zhou, "Developing ZigBee deployment guideline under WiFi interference for smart grid applications," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 110–120, Mar. 2011.

[14] C. Gezer and C. Buratti, "A ZigBee smart energy implementation for energy efficient buildings," in *Proc. IEEE 73rd Veh. Technol. Conf. (VTC Spring)*, May 15–18, 2011, pp. 1–5.

[15] R. P. Lewis, P. Igic, and Z. Zhongfu, "Assessment of communication methods for smart electricity metering in the U.K.," in *Proc. IEEE PES/IAS Conf. Sustainable Alternative Energy (SAE)*, Sep. 2009, pp. 1–4.

[16] A. Yarali, "Wireless mesh networking technology for commercial and industrial customers," in *Proc. Elect. Comput. Eng., CCECE*,May 1–4, 2008, pp. 000047–000052.

[17] M. Y. Zhai, "Transmission characteristics of low-voltage distribution networks in China under the smart grids environment," *IEEE Trans. Power Delivery*, vol. 26, no. 1, pp. 173–180, Jan. 2011.

[18] V. Paruchuri, A. Durresi, and M. Ramesh, "Securing powerline communications," in *Proc. IEEE Int. Symp. Power Line Commun. Appl., (ISPLC)*, Apr. 2–4, 2008, pp. 64–69.

[19] Q.Yang, J. A. Barria, and T. C. Green, "Communication infrastructures for distributed control of power distribution networks," *IEEE Trans. Ind. Inform.*, vol. 7, no. 2, pp. 316–327, May 2011.

[20] T. Sauter and M. Lobashov, "End-to-end communication architecture for smart grids," *IEEE Trans. Ind. Electron.*, vol. 58, no. 4, pp. 1218–1228, Apr. 2011.

[21] K. Moslehi and R. Kumar, “Smart grid—A reliability perspective,” *Innovative Smart Grid Technologies (ISGT)*, pp. 1–8, Jan. 19–21, 2010.

[22] Southern Company Services, Inc., “Comments request for information on smart grid communications requirements,” Jul. 2010

[23] R. Bo and F. Li, “Probabilistic LMP forecasting considering load uncertainty,” *IEEE Trans. Power Syst.*, vol. 24, pp. 1279–1289, Aug. 2009.

[24] *Power Line Communications*, H. Ferreira, L. Lampe, J. Newbury, and T. Swart (Editors), Eds. New York: Wiley, 2010.

[25] G. Bumiller, “Single frequency network technology for fast ad hoc communication networks over power lines,” WiKu-Wissenschaftsverlag Dr. Stein 2010.

[31] G. Bumiller, L. Lampe, and H. Hrasnica, “Power line communications for large-scale control and automation systems,” *IEEE Commun. Mag.*, vol. 48, no. 4, pp. 106–113, Apr. 2010.

[32] M. Biagi and L. Lampe, “Location assisted routing techniques for power line communication in smart grids,” in *Proc. IEEE Int. Conf. Smart Grid Commun.*, 2010, pp. 274–278.

[33] J. Sanchez, P. Ruiz, and R. Marin-Perez, “Beacon-less geographic routing made partical: Challenges, design guidelines and protocols,” *IEEE Commun. Mag.*, vol. 47, no. 8, pp. 85–91, Aug. 2009.

[34] N. Bressan, L. Bazzaco, N. Bui, P. Casari, L. Vangelista, and M. Zorzi, “The deployment of a smart monitoring system using wireless sensors and actuators networks,” in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, 2010, pp. 49–54.

[35] S. Dawson-Haggerty, A. Tavakoli, and D. Culler, “Hydro: A hybrid routing protocol for low-power and lossy networks,” in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, 2010, pp. 268–273.

[36] S. Goldfisher and S. J. Tanabe, “IEEE 1901 access system: An overview of its uniqueness and motivation,” *IEEE Commun. Mag.*, vol. 48, no. 10, pp. 150–157, Oct. 2010.

[37] V. C. Gungor, D. Sahin, T. Kocak, and S. Ergüt, “Smart grid communications and networking,” Türk Telekom, Tech. Rep. 11316-01, Apr 2011.