Statistics and Its Interface Volume 12 (2019) 467–478

The inverse gamma-difference distribution and its first moment in the Cauchy principal value sense

Aaron Hendrickson

In this paper, the probability density and distribution functions for the reciprocal-difference of independent gamma random variables with unequal shape parameters are derived. A theorem is developed and applied to evaluate the first moment of this distribution in the sense of the Cauchy principal value, which addresses the inverse chi-squared- and inverse exponential-difference distributions as special cases. These results are used to find the first moment and an approximation to the centralized inverse-Fano distribution, which models the sampling distribution of the photon transfer conversion gain measurement of electro-optical imaging sensors. A Monte Carlo simulation is performed to show how the first moment of the inverse gamma-difference distribution can be utilized to control the bias of the conversion gain measurement in a live experiment. The *low illumination problem* of conversion gain measurement is introduced with a discussion motivating future application of the theoretical results derived.

AMS 2000 subject classifications: Primary 30E20, 26A03, 78M99; secondary 62E17, 46N30, 33C20.
Keywords and phrases: Cauchy principal value, Centralized inverse-Fano distribution, Conversion gain, First negative moment, Gamma-difference distribution, Generalized central limit theorem, Hypergeometric function, Photon transfer.

1. INTRODUCTION

In Klar [1] a literature review was conducted to investigate an apparent lack of research done on the distribution of the difference of independent gamma random variables. It revealed that while this gamma-difference distribution seems largely absent from the statistical literature, it has in fact been studied for the case of equal shape parameters by several authors in a variety of different contexts [2, 3, 4, 5]. This, however, cannot be said for the unequal shape parameter case, for which only one explicit derivation by Mathai [6] as well as a brief mention in Krishna & Jose [7] was found. Sources for the distribution of the difference of independent chi-squared random variables with different degrees of freedom, which is a special case of the gamma-difference distribution with unequal shape parameters, were also provided [8, 9].

In a more recent article by Hendrickson [10], the gamma-difference distribution for unequal shape parameters was used in the derivation of the centralized inverse-Fano distribution, which was introduced as a model for the sampling distribution of the photon transfer conversion gain measurement of electro-optical imaging sensors [11]. A centralized inverse-Fano random variable was defined as

$$G = X/Y,$$

where $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y \sim \mathcal{GD}(\alpha_1, \alpha_2, \beta_1, \beta_2)$ are independent normal and gamma-difference random variables, respectively. Due to the functional complexity of this ratio, closed-form solutions for the density and distribution of G were only found for the case of integer shape parameters. For noninteger shape parameters, these functions were expressed analytically as integrals involving confluent hypergeometric functions which proved difficult to evaluate. It was suggested that when the conversion gain is measured under sufficiently high illumination, the dispersion of G is due almost entirely to the dispersion in Y [12, 10]; thus, opening up the possibility of approximating the density of G by a scaled inverse gamma-difference distribution. Furthermore, section 4 will show that various properties of G can be drawn out from a characterization of the inverse gamma-difference distribution including, in particular, its first moment. Understanding these properties is of interest in the imaging science community since a better understanding of the behavior of the conversion gain measurement, especially under low illumination conditions, is critical to properly characterizing the performance of sensors with nonlinear transfer functions as well as developing a generalized photon transfer methodology.

In light of these observations, this paper seeks to provide an initial characterization of the distribution representing the reciprocal difference of independent gamma random variables with unequal shape parameters. This will be accomplished by deriving some preliminary results pertaining to the gamma-difference distribution followed by a derivation of the inverse distribution and its first moment using the Cauchy principal value interpretation. This paper concludes with a discussion on an application of these theoretical results involving electro-optical imaging sensors.

2. GAMMA-DIFFERENCE DISTRIBUTIONS

The starting point for this characterization is uncovering the pertinent details of the gamma-difference distribution as well as deriving some new results for subsequent use in section 3. The gamma-difference random variable is

$$Y = Y_1 - Y_2,$$

where $Y_1 \sim \mathcal{G}(\alpha_1, \beta_1)$ and $Y_2 \sim \mathcal{G}(\alpha_2, \beta_2)$ are independent gamma random variables parameterized in terms of a shape parameter $\alpha_i > 0$ and rate parameter $\beta_i > 0$ for $i = 1, 2$. For the density and distribution functions of Y^{-1} in section 3.1 three results are needed, namely, the density and distribution of Y as well as $\mathrm{pr}(Y \le 0)$. Then, for the derivation of the principal-valued expectation of Y^{-1} in section 3.2 we require the density of Y at the origin as well an expression for the noninteger moments of Y partitioned about the origin[1]. Here we present all of these results albeit not in this order.

The density of Y, as presented in [6] and [1], is expressed in terms of Whittaker's confluent hypergeometric function $W_{\kappa,\mu}(z)$. Using the relation found in [13, Eq. 13.14.3], which expresses $W_{\kappa,\mu}(z)$ in terms of Kummer's confluent hypergeometric function of the 2nd kind, $U(a, b, z)$, the density of Y can be expressed in a more compact form.

$$(1)\quad f_Y(y) = \begin{cases} \frac{C_Y}{\Gamma(\alpha_2)}\, e^{\beta_2 y}\, U(1-\alpha_2, 2-\alpha_o, -\beta_o y), & y \le 0, \\ \frac{C_Y}{\Gamma(\alpha_1)}\, e^{-\beta_1 y}\, U(1-\alpha_1, 2-\alpha_o, \beta_o y), & y > 0, \end{cases}$$

where $\alpha_o = \alpha_1 + \alpha_2$, $\beta_o = \beta_1 + \beta_2$, $C_Y = \beta_1^{\alpha_1} \beta_2^{\alpha_2} \beta_o^{1-\alpha_o}$, and $\Gamma(s)$ is the gamma function. An expression for the density of Y at the origin is now provided in the following.

Lemma 2.1. *For α_o and β_o as defined following (1) and $Y \sim \mathcal{GD}(\alpha_1, \alpha_2, \beta_1, \beta_2)$*

$$f_Y(0) = \begin{cases} \infty, & 0 < \alpha_o \le 1, \\ \frac{\beta_1^{\alpha_1} \beta_2^{\alpha_2} \beta_o^{1-\alpha_o}}{(\alpha_o - 1)\mathrm{B}(\alpha_1, \alpha_2)}, & \alpha_o > 1, \end{cases}$$

where $\mathrm{B}(x, y) := \Gamma(x)\Gamma(y)/\Gamma(x + y)$ is the beta function.

Proof. From (1), $f_Y(0)$ is equivalent to

$$f_Y(0) = \frac{C_Y}{\Gamma(\alpha_i)} \lim_{y \to 0^+} U(1 - \alpha_i, 2 - \alpha_o, \beta_o y).$$

Let $a = 1-\alpha_i < 1$ and $b = 2-\alpha_o < 2$, then the limiting cases of $U(a, b, z)$ as $z \to 0$ can be found in [13, Eqs. 13.2.18–22].

$$\lim_{z \to 0} U(a, b, z) = \begin{cases} \infty, & 1 \le b < 2, \\ \frac{\Gamma(1-b)}{\Gamma(a-b+1)}, & b < 1. \end{cases}$$

Substituting in the proper values for a and b and multiplying by $C_Y/\Gamma(\alpha_i)$ yields the final result. □

As for the distribution of Y, there is no known closed-form expression so it is written as follows [1].

$$(2)\quad F_Y(y) = \frac{\beta_2^{\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_{\max\{0,-y\}}^{\infty} t^{\alpha_2 - 1} e^{-\beta_2 t} \gamma\left(\alpha_1, \beta_1(t + y)\right) \mathrm{d}t,$$

where $\gamma(s, z) := \int_0^z t^{s-1}\, e^{-t}\, \mathrm{d}t$ is the lower-incomplete gamma function. To evaluate $\mathrm{pr}(Y \le 0)$ it will be easier to first find the needed expression for the partitioned noninteger moments of Y and then recover the desired probability from the zeroth moment. To obtain the partitioned moments we start with decomposing the gamma-difference variable into its positive and negative components by $Y = Y\mathbb{1}_{Y\le 0} + Y\mathbb{1}_{Y>0}$, where $\mathbb{1}_A$ is the indicator function on the set A, and then introduce the following lemma.

Lemma 2.2 (13, Eq. 13.10.7)**.** *For $\Re\nu > \max\{\Re b - 1, 0\}$ and $\Re z > 0$*

$$\int_0^{\infty} t^{\nu-1} e^{-zt}\, U(a, b, t)\, \mathrm{d}t = \frac{\Gamma(\nu)\Gamma(\nu - b + 1)}{\Gamma(a - b + 1 + \nu)} z^{-\nu} {}_2F_1\left(\begin{matrix} a, \nu \\ a - b + 1 + \nu \end{matrix}; 1 - \tfrac{1}{z}\right),$$

where

$$(3)\qquad {}_pF_q\left(\begin{matrix} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{matrix}; z\right) := \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{z^n}{n!}$$

is the generalized hypergeometric function, and $(s)_n := \Gamma(s+n)/\Gamma(s)$ is the rising Pochhammer symbol.

Lemma 2.2 is then applied to evaluate $E(Y^{\varepsilon-1}\mathbb{1}_{Y\le 0})$ and $E(Y^{\varepsilon-1}\mathbb{1}_{Y>0})$. For the negative component, i.e. $Y\mathbb{1}_{Y\le 0}$, we define the branch cut $z^\varepsilon = |z|^\varepsilon e^{\mathrm{i}\varepsilon \arg z}$ for $-\pi/2 < \arg z < 3\pi/2$ and make the substitution $y = -t$ to put the integral defining $E(Y^{\varepsilon-1}\mathbb{1}_{Y\le 0})$ into a useful form.

Lemma 2.3. *For α_o and C_Y as defined following (1) and $Y \sim \mathcal{GD}(\alpha_1, \alpha_2, \beta_1, \beta_2)$ let $\varepsilon > \max\{1 - \alpha_o, 0\}$, then $E(Y^{\varepsilon-1}) = E(Y^{\varepsilon-1}\mathbb{1}_{Y\le 0}) + E(Y^{\varepsilon-1}\mathbb{1}_{Y>0})$ where*

$$E(Y^{\varepsilon-1}\mathbb{1}_{Y\le 0}) = -e^{\mathrm{i}\pi\varepsilon} C_Y \frac{\Gamma(\varepsilon)\Gamma(\alpha_o - 1 + \varepsilon)}{\Gamma(\varepsilon + \alpha_1)\Gamma(\alpha_2)\beta_2^{\varepsilon}}\, {}_2F_1\left(\begin{matrix} 1 - \alpha_2, \varepsilon \\ \varepsilon + \alpha_1 \end{matrix}; -\tfrac{\beta_1}{\beta_2}\right),$$

$$E(Y^{\varepsilon-1}\mathbb{1}_{Y>0}) = C_Y \frac{\Gamma(\varepsilon)\Gamma(\alpha_o - 1 + \varepsilon)}{\Gamma(\alpha_1)\Gamma(\varepsilon + \alpha_2)\beta_1^{\varepsilon}}\, {}_2F_1\left(\begin{matrix} 1 - \alpha_1, \varepsilon \\ \varepsilon + \alpha_2 \end{matrix}; -\tfrac{\beta_2}{\beta_1}\right).$$

The partitioned noninteger moments can now be utilized to find the probability of Y being nonpositive as well as the complementary expression by substituting $\varepsilon = 1$ into each result of Lemma 2.3.

[1] By this we mean the integral defining the moments is partitioned about the origin.

Corollary 2.1. *For α_o and β_o as defined following (1) and $Y \sim \mathcal{GD}(\alpha_1, \alpha_2, \beta_1, \beta_2)$*

$$\mathrm{pr}(Y \le 0) = \frac{\beta_1^{\alpha_1}\beta_2^{\alpha_2-1}\beta_o^{1-\alpha_o}}{\alpha_1\,\mathrm{B}(\alpha_1,\alpha_2)}\,{}_2F_1\left(\begin{matrix}1-\alpha_2,\,1\\ 1+\alpha_1\end{matrix};-\tfrac{\beta_1}{\beta_2}\right),$$
$$\mathrm{pr}(Y > 0) = \frac{\beta_1^{\alpha_1-1}\beta_2^{\alpha_2}\beta_o^{1-\alpha_o}}{\alpha_2\,\mathrm{B}(\alpha_1,\alpha_2)}\,{}_2F_1\left(\begin{matrix}1-\alpha_1,\,1\\ 1+\alpha_2\end{matrix};-\tfrac{\beta_2}{\beta_1}\right).$$

3. THE INVERSE GAMMA-DIFFERENCE DISTRIBUTION

In this section three results are derived in the form of the density and distribution functions of the inverse gamma-difference variable as well as its principal-valued first moment. In regards to the latter, Theorem 3.2 and Lemmas 3.1–3.2 will be developed to facilitate the evaluation of the first moment found in the main result of Theorem 3.3. Along the way several additional auxiliary result will also be presented.

3.1 Density and distribution functions

With the requisite description of the gamma-difference distribution at hand, the inverse gamma-difference variable $Z = Y^{-1}$ is introduced. The density and distribution of Z are presented in Theorem 3.1.

Theorem 3.1. *For $Y \sim \mathcal{GD}(\alpha_1, \alpha_2, \beta_1, \beta_2)$ let $Z = Y^{-1}$, then*

$$(i)\quad f_Z(z) = \begin{cases} z^{-2}\, f_Y(z^{-1}), & z \neq 0,\\ 0, & z = 0,\end{cases}$$

$$(ii)\quad F_Z(z) = \mathbb{1}_{(0,\infty)}(z) + F_Y(0) - \begin{cases} F_Y(z^{-1}), & z \neq 0,\\ 0, & z = 0,\end{cases}$$

where f_Y is defined in (1), F_Y is defined in (2), and $F_Y(0) = \mathrm{pr}(Y \le 0)$ is defined in Corollary 2.1.

Proof (i). Z is a univariate transformation of Y; thus, the change of variables formula for density functions is applied to derive the $z \neq 0$ case of f_Z directly from the density of Y in (1). For the $z = 0$ case we need to evaluate the limit of f_Z as $z \to 0^-$ or $z \to 0^+$. Taking the limit as $z \to 0^+$ and substituting $y = 1/z$ yields $f_Z(0) = \lim_{y\to\infty} y^2 f_Y(y)$. As $y \to \infty$, $U(a, b, y) \sim y^{-a}$ [13, Eq. 13.2.6]; therefore, $f_Z(0) \propto \lim_{y\to\infty} y^{\alpha_1-1}e^{-\beta_1 y} = 0$. □

Proof (ii). A straightforward integration of f_Z yields F_Z. □

3.2 Moments

Since the inverse gamma-difference variable Z is the reciprocal transformation of the gamma-difference variable Y, the moments $E(Z^n)$ are equivalent to the negative moments $E(Y^{-n})$.

Proposition 3.1. *For α_o as defined following (1) and $Z \sim \mathcal{IGD}(\alpha_1, \alpha_2, \beta_1, \beta_2)$, $E(Z^n)$ exists and is finite for $n < \min\{\alpha_o, 1\}$.*

Proof. By definition, $E(Z^n) = E(Y^{-n})$. Let $n = 1 - \varepsilon$, then according to Lemma 2.3, $E(Y^{-n})$ exists and is finite for $1 - n < \max\{1 - \alpha_o, 0\}$. Solving this inequality for n yields $n < \min\{\alpha_o, 1\}$. □

Proposition 3.1 indicates that Z does not admit a first moment. This observation is confirmed by the work of Piegorsch & Casella which demonstrated that the reciprocal of a continuous random variable X does not possess a defined expectation if $f_X(0) > 0$ [14]. From Lemma 2.1 the density of Y is shown to always be nonzero at the origin; therefore, we employ the concept of the Cauchy principal value to define the *principal-valued first negative moment* as an alternative interpretation to the traditional definition.

Definition 3.1. *Let $\delta_0, \delta_1 > 0$ and X be a continuous random variable with density f_X defined on $\mathbb{R}$. The principal-valued first negative moment of X is*

$$PVE(X^{-1}) := \lim_{\delta_0,\delta_1\to 0}\left(\int_{-1/\delta_1}^{-\delta_0} + \int_{\delta_0}^{1/\delta_1}\right)\frac{f_X(x)}{x}\,\mathrm{d}x,$$

if and only if each integral exists and is finite.

The use of the Cauchy principal value prescription for first negative moments has most notably been applied to problems arising from reciprocals of normal, skew-normal, skew-t, and generalized Student-t random variables [15, 16, 17]. In the context of the generalized central limit theorem, the principal-valued first negative moment also appears as a centering constant for describing the limiting distribution of suitably normed sums of reciprocal random variables. Let $\{X_i\}$ be a sequence of i.i.d. random variables with density f_X that is continuous and nonzero at the origin, then according to the generalized central limit theorem, X^{-1} belongs to the domain of attraction of the Cauchy law for which we write [18]

$$\frac{\frac{1}{n}\sum_{i=1}^{n} X_i^{-1} - PVE(X^{-1})}{\pi f_X(0)} \xrightarrow{d} \mathcal{C}(0,1),$$

where $\mathcal{C}(0,1)$ represents the standard Cauchy distribution.

While several applications exist, evaluating principal-valued first negative moments for certain distributions via Definition 3.1 may not prove to be tractable motivating the need for other approaches. The following theorem presents one such alternative approach.

Theorem 3.2. *Let X be a continuous random variable with density f_X where $f_X(x) > 0$ over some open interval $\Omega \subset \mathbb{R}$ containing the origin. If f_X is continuously differentiable on Ω and $|\frac{\mathrm{d}}{\mathrm{d}x} f_X(x)| < M$ holds in a neighborhood of the origin for $M \in \mathbb{R}$ then for $\varepsilon > 0$*

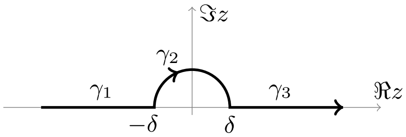

Figure 1. The contour $\mathcal{C}$.

(4) $$PVE(X^{-1}) = \lim_{\epsilon\to 0}\left[E(X^{\epsilon-1}\mathbb{1}_{Y\le 0}) + E(X^{\epsilon-1}\mathbb{1}_{Y>0})\right] + f_X(0)\pi\mathrm{i}.$$

Proof. For notational convenience we let $\delta > 0$ and define

$$X_\delta = X\mathbb{1}_{X\notin[-\delta,\delta]} + X\mathbb{1}_{X\in[-\delta,\delta]}.$$

Without loss of generality, assume $f_X(x) > 0$ for all $x \in \mathbb{R}$ and let $\varepsilon > 0$, then

(5) $$E(X_\delta^{\varepsilon-1}) = E(X^{\varepsilon-1}\mathbb{1}_{X\notin[-\delta,\delta]}) + E(X^{\varepsilon-1}\mathbb{1}_{X\in[-\delta,\delta]}).$$

Changing the value of δ only serves to change the partition of the sample space of X. As such, it should be readily apparent that $E(X^n) = E(X_\delta^n)$ for any δ so long as $E(X^n)$ exists. Now, for $z \in \mathbb{C}$ we define the branch cut $z^\varepsilon = |z|^\varepsilon e^{\mathrm{i}\varepsilon \arg z}$ with $-\pi/2 < \arg z < 3\pi/2$ and then consider the contour $\mathcal{C}$ depicted in Figure 1. If for some δ the density f_X is analytic over the region bound by the real axis and the contour γ_2, that is, $R = \{z : |z| \le \delta \wedge \Im z \ge 0\}$, then the expected value in (5) is equal to the integral along $\mathcal{C}$. For the integral over γ_1 and γ_3 we have

$$\lim_{\varepsilon\to 0}\int_{|x|>\delta} x^\varepsilon \frac{f_X(x)}{x}\,\mathrm{d}x = \int_{|x|>\delta}\frac{f_X(x)}{x}\,\mathrm{d}x,$$

by the dominated convergence theorem. Furthermore, for the integral over γ_2 we obtain the result

$$\lim_{\varepsilon\to 0}\mathrm{i}\delta^\varepsilon\int_\pi^0 e^{\mathrm{i}\varepsilon\varphi} f_X(\delta e^{\mathrm{i}\varphi})\,\mathrm{d}\varphi = \mathrm{i}\int_\pi^0 f_X(\delta e^{\mathrm{i}\varphi})\,\mathrm{d}\varphi,$$

since $\lim_{\varepsilon\to 0} e^{\mathrm{i}\varepsilon\varphi} \to 1$ uniformly for $\varphi \in [0,\pi]$. Bringing both terms together then provides the limiting expression

(6) $$\lim_{\varepsilon\to 0} E(X_\delta^{\varepsilon-1}) = \int_{|x|>\delta}\frac{f_X(x)}{x}\,\mathrm{d}x + \mathrm{i}\int_\pi^0 f_X(\delta e^{\mathrm{i}\varphi})\,\mathrm{d}\varphi.$$

Under the stated assumptions, (6) is independent of the choice for δ. Therefore, taking $\delta \to 0$ yields

(7) $$\lim_{\varepsilon\to 0} E(X_0^{\varepsilon-1}) = \mathrm{PV}\int_{\mathbb{R}}\frac{f_X(x)}{x}\,\mathrm{d}x - f_X(0)\pi\mathrm{i}.$$

The work by Peng (2008) showed that if f_X is continuously differentiable on Ω and $|\frac{\mathrm{d}}{\mathrm{d}x}f_X(x)| < M$ holds in a neighborhood of the origin for $M \in \mathbb{R}$, then this limit exists and is finite [16]. Upon adding $f_X(0)\pi\mathrm{i}$ to both sides and noting that $E(X_0^{\varepsilon-1}) = E(X^{\epsilon-1}\mathbb{1}_{Y\le 0}) + E(X^{\epsilon-1}\mathbb{1}_{Y>0})$, the expression in (4) is attained. □

The last two results needed to evaluate the principal-valued first moment of the inverse gamma-difference distribution, $PVE(Z)$, are provided in the following two lemmas. The derivations of these results are helpful, although not necessary, to understand the steps taken in the final evaluation of $PVE(Z)$ presented in Theorem 3.3.

Lemma 3.1. *Let* $f(\varepsilon) = {}_2F_1(1-a, \varepsilon;\, \varepsilon + b;\, -z)$ *where* $a, b, z, \varepsilon > 0$, $a + b > 1$, *and* $f' = \mathrm{d}f/\mathrm{d}\varepsilon$, *then*

(*i*) $f(0) = 1$,

(*ii*) $$f'(0) = \frac{(a-1)\,z}{b}\,{}_3F_2\left(\begin{matrix}2-a,\,1,\,1\\ 1+b,\,2\end{matrix};\,-z\right).$$

Proof (*i*)*.* Provided the definition of the Pochhammer symbol and the generalized hypergeometric function in Lemma 2.2, one finds $(0)_{k>0} = 0$ and ${}_2F_1(1-a, 0; b; -z) = 1$, respectively. □

Proof (*ii*)*.* Making use of the generalized hypergeometric series in Lemma 2.2, f is expressed as

(8) $$f(\varepsilon) = \sum_{k=0}^{\infty}\xi_k(\varepsilon)\,(1-a)_k\frac{(-z)^k}{k!},$$

where $\xi_k(\varepsilon) = (\varepsilon)_k/(\varepsilon+b)_k$. Assume $a + b > 1$ and $z \le 1$, then the series in (8) converges absolutely and

(9) $$f'(\varepsilon) = \sum_{k=1}^{\infty}\xi_k'(\varepsilon)\,(1-a)_k\frac{(-z)^k}{k!}.$$

To evaluate the derivative, we introduce the differential formula $\mathrm{d}(s)_n/\mathrm{d}s = (s)_n(\psi(s+n) - \psi(s))$ where $\psi(s) := \frac{\mathrm{d}}{\mathrm{d}s}\log\Gamma(s)$ is the digamma function [19, Eq. 3.3]. The quantity $\xi_k'(\varepsilon)$ is subsequently found to be

$$\xi_k'(\varepsilon) = \frac{(\varepsilon)_k}{(\varepsilon+b)_k}\Big(\psi(\varepsilon+b) - \psi(\varepsilon+b+k) + \psi(\varepsilon+k) - \psi(\varepsilon)\Big).$$

Now writing $(\varepsilon)_k$ in terms of gamma functions and making use of of the recurrence relations $\Gamma(\varepsilon) = \Gamma(\varepsilon+1)/\varepsilon$ and $\psi(\varepsilon) = \psi(\varepsilon+1) - \varepsilon^{-1}$ we write

$$\xi_k'(\varepsilon) = \frac{\Gamma(\varepsilon+k)}{(\varepsilon+b)_k\Gamma(\varepsilon+1)}\Big(1 + \varepsilon\big[\psi(\varepsilon+b) - \psi(\varepsilon+b+k) + \psi(\varepsilon+k) - \psi(\varepsilon+1)\big]\Big).$$

From this expression we find $\xi_k'(0) = \Gamma(k)/(b)_k$; thus,

$$f'(0) = -z\sum_{k=0}^{\infty}\frac{(1-a)_{k+1}\Gamma(k+1)}{(b)_{k+1}(k+1)}\frac{(-z)^k}{k!}.$$

By the properties of the Pochhammer symbol $\Gamma(k+1) = (1)_k$, $k + 1 = (2)_k/(1)_k$, and $(s)_{k+1} = s(1+s)_k$, which results in the series taking on the form

$$f'(0) = \frac{(a-1)\,z}{b} \sum_{k=0}^{\infty} \frac{(2-a)_k (1)_k (1)_k}{(1+b)_k (2)_k} \frac{(-z)^k}{k!}. \tag{10}$$

By inspection of the generalized hypergeometric series in Lemma 2.2, the desired result is realized. □

Remark 3.1. *For $z > 1$ the result in (10) is defined by the analytic continuation of the generalized hypergeometric function ${}_3F_2(\mathbf{a};\mathbf{b};z)$.*

Lemma 3.2. *Let $g(\varepsilon) = \Gamma(a)\Gamma(\varepsilon+b)c^{\varepsilon}$ where $a, b, c, \varepsilon > 0$ and $g' = \mathrm{d}g/\mathrm{d}\varepsilon$, then*

(i) $g(0) = \Gamma(a)\Gamma(b)$,

(ii) $g'(0) = \Gamma(a)\Gamma(b)(\psi(b) + \log c)$.

Theorem 3.3. *For α_o and β_o as defined following (1) and $Z \sim \mathcal{IGD}(\alpha_1, \alpha_2, \beta_1, \beta_2)$, let $\alpha_o > 1$, then*

$$PVE(Z) = \frac{\beta_1^{\alpha_1}\beta_2^{\alpha_2}\beta_o^{1-\alpha_o}}{(\alpha_o-1)\mathrm{B}(\alpha_1,\alpha_2)} \Bigg(\psi(\alpha_1) - \log\beta_1 + \frac{(\alpha_1-1)\,\beta_2}{\alpha_2\,\beta_1} {}_3F_2\left(\begin{matrix} 2-\alpha_1,\,1,\,1 \\ 1+\alpha_2,\,2 \end{matrix} ; -\tfrac{\beta_2}{\beta_1}\right) - \psi(\alpha_2) + \log\beta_2 - \frac{(\alpha_2-1)\,\beta_1}{\alpha_1\,\beta_2} {}_3F_2\left(\begin{matrix} 2-\alpha_2,\,1,\,1 \\ 1+\alpha_1,\,2 \end{matrix} ; -\tfrac{\beta_1}{\beta_2}\right)\Bigg).$$

Proof. Let $\varepsilon > 0$. From Theorem 3.2 we note that $\lim_{\varepsilon\to 0} E(Y_0^{\varepsilon-1})$ is comprised of a principal-valued contribution that is real-valued, as well as a purely imaginary term that is the residue from integrating over the simple pole at the origin for $\varepsilon = 0$. With this in mind, the principal-valued first moment of Z is

$$PVE(Z) = \Re \lim_{\epsilon\to 0} \left[E(X^{\epsilon-1}\mathbb{1}_{Y\le 0}) + E(X^{\epsilon-1}\mathbb{1}_{Y>0}) \right].$$

Assume $\alpha_o > 1$ such that f_Y is finite at the origin, then by Lemma 2.3

$$\lim_{\varepsilon\to 0} E(Y_0^{\varepsilon-1}) = C_Y \Gamma(\alpha_o - 1) \lim_{\varepsilon\to 0} \Gamma(\varepsilon)\, D(\varepsilon),$$

where

$$D(\varepsilon) = \frac{{}_2F_1\left(\begin{smallmatrix} 1-\alpha_1,\,\varepsilon \\ \varepsilon+\alpha_2 \end{smallmatrix} ; -\frac{\beta_2}{\beta_1}\right)}{\Gamma(\alpha_1)\Gamma(\varepsilon+\alpha_2)\beta_1^{\varepsilon}} - e^{\mathrm{i}\pi\varepsilon} \frac{{}_2F_1\left(\begin{smallmatrix} 1-\alpha_2,\,\varepsilon \\ \varepsilon+\alpha_1 \end{smallmatrix} ; -\frac{\beta_1}{\beta_2}\right)}{\Gamma(\varepsilon+\alpha_1)\Gamma(\alpha_2)\beta_2^{\varepsilon}}.$$

From the definition of the generalized hypergeometric function, ${}_2F_1(a, 0; c; z) = 1$; thus, it is straightforward to show $D(0) = 0$. Furthermore, by writing the leading gamma term as $\Gamma(\varepsilon) = \Gamma(\varepsilon+1)/\varepsilon$ it becomes obvious that $\lim_{\varepsilon\to 0} E(Y_0^{\varepsilon-1}) \propto \lim_{\varepsilon\to 0} D(\varepsilon)/\varepsilon$, which is indeterminate. By L'Hôpital's rule

$$\lim_{\varepsilon\to 0} E(Y_0^{\varepsilon-1}) = \frac{C_Y\Gamma(\alpha_o)}{(\alpha_o-1)} \lim_{\varepsilon\to 0} \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \left(\frac{f_1(\varepsilon)}{g_1(\varepsilon)} - e^{\mathrm{i}\pi\varepsilon} \frac{f_2(\varepsilon)}{g_2(\varepsilon)} \right), \tag{11}$$

where

$$f_1(\varepsilon) = {}_2F_1\left(\begin{matrix} 1-\alpha_1,\,\varepsilon \\ \varepsilon+\alpha_2 \end{matrix} ; -\tfrac{\beta_2}{\beta_1}\right),$$
$$f_2(\varepsilon) = {}_2F_1\left(\begin{matrix} 1-\alpha_2,\,\varepsilon \\ \varepsilon+\alpha_1 \end{matrix} ; -\tfrac{\beta_1}{\beta_2}\right),$$

and

$$g_1(\varepsilon) = \Gamma(\alpha_1)\Gamma(\varepsilon+\alpha_2)\beta_1^{\varepsilon},$$
$$g_2(\varepsilon) = \Gamma(\varepsilon+\alpha_1)\Gamma(\alpha_2)\beta_2^{\varepsilon}.$$

Evaluating the derivative in (11) and making use of Lemma 2.1 gives

$$\lim_{\varepsilon\to 0} E(Y_0^{\varepsilon-1}) = f_Y(0)\Bigg(f_1'(0) - \frac{f_1(0)\, g_1'(0)}{g_1(0)} - f_2'(0) + \frac{f_2(0)\, g_2'(0)}{g_2(0)} \Bigg) - f_Y(0)\pi\mathrm{i}. \tag{12}$$

Lemmas 2.1, 3.1, and 3.2 provide the results necessary to evaluate the individual components. After a bit of algebraic manipulation the real part of (12) yields the desired expression for $PVE(Z)$. □

As was the case for the noninteger moments of Y presented in Lemma 2.3, $PVE(Z)$ consists of the sum of two terms which are the contributions of the negative and positive cases of f_Z. This property in turn allows for approximations when Z is nearly positive or negative.

Corollary 3.1. *In Theorem 3.3, as $\mathrm{pr}(Z > 0) \to 1$*

$$PVE(Z) \to \frac{\beta_1^{\alpha_1}\beta_2^{\alpha_2}\beta_o^{1-\alpha_o}}{(\alpha_o-1)\mathrm{B}(\alpha_1,\alpha_2)} \Bigg(\psi(\alpha_1) - \log\beta_1 + \frac{(\alpha_1-1)\,\beta_2}{\alpha_2\,\beta_1} {}_3F_2\left(\begin{matrix} 2-\alpha_1,\,1,\,1 \\ 1+\alpha_2,\,2 \end{matrix} ; -\tfrac{\beta_2}{\beta_1}\right)\Bigg).$$

Likewise, as $\mathrm{pr}(Z < 0) \to 1$

$$PVE(Z) \to -\frac{\beta_1^{\alpha_1}\beta_2^{\alpha_2}\beta_o^{1-\alpha_o}}{(\alpha_o-1)\mathrm{B}(\alpha_1,\alpha_2)} \Bigg(\psi(\alpha_2) - \log\beta_2 + \frac{(\alpha_2-1)\,\beta_1}{\alpha_1\,\beta_2} {}_3F_2\left(\begin{matrix} 2-\alpha_2,\,1,\,1 \\ 1+\alpha_1,\,2 \end{matrix} ; -\tfrac{\beta_1}{\beta_2}\right)\Bigg).$$

In certain applications involving gamma-difference distributions with large shape parameters (α_i), it can be beneficial to approximate $PVE(Z)$ via the normal approximation. For a derivation of the normal approximation using Theorem 3.2 see appendix B.

Corollary 3.2. *For $Z \sim \mathcal{IGD}(\alpha_1, \alpha_2, \beta_1, \beta_2)$ and large α_1 and α_2*

$$PVE(Z) \approx \frac{\sqrt{2}}{\sqrt{\alpha_1/\beta_1^2 + \alpha_2/\beta_2^2}} \mathcal{D}\left(\frac{\alpha_1/\beta_1 - \alpha_2/\beta_2}{\sqrt{2(\alpha_1/\beta_1^2 + \alpha_2/\beta_2^2)}} \right),$$

where $\mathcal{D}(z) := e^{-z^2}\int_0^z e^{t^2}\,\mathrm{d}t$ is the Dawson integral.

Additionally, it is well known that the chi-squared and exponential distributions are special cases of the gamma distribution. As a consequence of this, $PVE(Z)$ admits the principal-valued first moment of the reciprocal difference of independent chi-squared and exponential random variables as special cases.

Corollary 3.3. *For $i = 1, 2$ and $\nu_i > 0$, let $Y_i \sim \chi^2(\nu_i) = \mathcal{G}(\nu_i/2, 2)$, $\nu_o = \nu_1 + \nu_2$, and $Z = (Y_1 - Y_2)^{-1}$, then for $\nu_o > 2$*

$$PVE(Z) = \frac{2^{2-\nu_o/2}}{\left(\frac{\nu_o}{2} - 1\right) \mathrm{B}\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)} \Bigg(\psi\left(\frac{\nu_1}{2}\right) + \frac{\nu_1 - 2}{\nu_2} \, {}_3F_2\left(\begin{matrix} 2 - \frac{\nu_1}{2}, 1, 1 \\ 1 + \frac{\nu_2}{2}, 2 \end{matrix}; -1\right) - \psi\left(\frac{\nu_2}{2}\right) - \frac{\nu_2 - 2}{\nu_1} \, {}_3F_2\left(\begin{matrix} 2 - \frac{\nu_2}{2}, 1, 1 \\ 1 + \frac{\nu_1}{2}, 2 \end{matrix}; -1\right) \Bigg),$$

Corollary 3.4. *For $i = 1, 2$ and $\lambda_i > 0$, let $Y_i \sim \mathrm{Exp}(\lambda_i) = \mathcal{G}(1, \lambda_i)$ and $Z = (Y_1 - Y_2)^{-1}$, then*

$$PVE(Z) = \frac{\log(\lambda_2/\lambda_1)}{\lambda_1^{-1} + \lambda_2^{-1}}.$$

4. APPLICATION

4.1 Preliminaries

Photon transfer (PT) is a methodology for characterizing the performance of electro-optical imaging sensors. Due in part to its simplicity and broad applicability to a variety of sensor technologies, PT has been adopted industry-wide as the archetypal approach to measuring sensor performance since its conception in the mid 1970s. For the manufacturers of imaging sensors, PT is used in the design and optimization processes to ensure the best possible product is created. On the consumer end, PT is used to verify performance claims made by sensor manufacturers and to calibrate imaging systems for use in a specific application. As a consequence of this widespread use, the PT method was officially standardized in the European machine vision association's (EMVA) 1288 standard for sensor evaluation which had its first release in 2005.

The primary purpose of PT is measuring a sensor's transfer function, that is, the parametric relationship between the input signal, in the form of packets of electrons (e-) produced in the pixels by interacting photons, to the output signal usually in the form of digital numbers (DN). When this relationship is linear, the transfer function can be described by a single constant known as the conversion gain g in units of (e-/DN). Measuring the conversion gain is central to the photon transfer method as it effectively allows one to convert the output signal in arbitrary digital units into a physically meaningful quantity of electrons. This in turn allows for the pixels comprising the sensor to be characterized in terms of key imaging performance metrics such as read noise, well-capacity, dynamic range, and dark current in an absolute sense [11]. Consider for example, the read noise which dictates the lowest light conditions under which a sensor can generate useful imagery. The read noise is measured according to the formula

$$\sigma_r(e\text{-}) = g(e\text{-}/\mathrm{DN}) \times \sigma_r(\mathrm{DN}),$$

where $\sigma_r(\mathrm{DN})$ is the standard deviation of the digital output signal of a pixel when exposed to zero illumination at a zero second integration time and $\sigma_r(e\text{-})$ is the read noise in electron units. Like the read noise, each of the other performance metrics mentioned, as well as others not mentioned, require multiplying the gain by some sample statistic in DN units to convert it to electron units. As a result of the central role it plays, the uncertainty and bias of the conversion gain measurement fundamentally limits the uncertainty and bias the PT methodology as a whole. That said, little is actually understood about the properties of the conversion gain measurement; thus, motivating the need for a detailed analysis.

The earliest work carried out to characterize this fundamental measurement traces back to that presented in [11, 12] where estimates for the uncertainty were obtained via propagation of errors. These results provided a reasonable approximation to the uncertainty of the measurement under limited circumstances. In particular, both works assume the gain is measured under high illumination conditions to achieve a so-called *shot noise-limited* response. While adequate in many applications, this assumption is much too restrictive for the characterization of certain sensor architectures as discussed in more detail in section 4.5. In response to this limitation, the centralized inverse-Fano (CIF) distribution was derived as a model of the sampling distribution for the PT conversion gain measurement [10]. With a theoretical model of the sampling distribution the properties of the measurement and ultimately how it behaves under low illumination conditions could begin to be explored and subsequently used to develop a much needed approach to low-illumination conversion gain measurement. Such an approach would be a critical component to the developing a more generalized PT methodology.

4.2 An estimator for g

Developing an estimator for the conversion gain of a linear pixel is done by taking advantage of the statistical properties of photons. As a consequence of Bose-Einstein statistics, photons–which we will denote with the unit symbol γ–can be modeled as Poisson random variables whereby the number of photons k observed by a pixel over some integration time τ (s) is described by the Poisson probability mass function

$$p(k) = e^{-\Phi\tau} \frac{(\Phi\tau)^k}{k!}, \quad k = 0, 1, 2, \ldots,$$

where Φ is the photon flux with units ($\gamma \cdot \mathrm{px}^{-1} \cdot \mathrm{s}^{-1}$), [11, 12, 20]. Noting that the Poisson mass function has equal mean and variance and assuming the pixel produces no signal and noise of its own, the conversion gain is the mean-to-variance ratio of the pixel's digital output. However, because the pixel will produce its own background signal and noise, the conversion gain can be measured according to the following formula [11, 10].

$$G = (\bar{X} - \bar{Y})/(\hat{X} - \hat{Y}) = \bar{P}/\hat{P}$$

The statistics $\bar{X} = \sum_{i=1}^{n_1} X_i/n_1$ and $\hat{X} = \sum_{i=1}^{n_1}(X_i - \bar{X})^2/(n_1 - 1)$ are the sample mean and variance computed from n_1 i.i.d. digital observations of a pixel under illumination. Similarly, the statistics $\bar{Y}$ and $\hat{Y}$ are the sample mean and variance computed from n_2 i.i.d. digital observations of the same pixel under zero illumination. Thus, the differences $\bar{P}$ and $\hat{P}$ are estimators for the photon induced signal, $E(\bar{P}) > 0$, and signal variance, $E(\hat{P}) > 0$, and their ratio estimates the desired estimand $g = E(\bar{P})\,E(\hat{P})^{-1}$ which is the gain. According to [10, 11, 12] the total pixel noise under illumination and non-illumination can be adequately modeled as normal such that $X_i \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $Y_i \sim \mathcal{N}(\mu_2, \sigma_2^2)$. In short, the justification for this model is tied to the sum of the various noise sources present in the sensor being normally distributed while the Poisson distribution of the photon noise approaches a normal form for even moderately small photon exposures. This model inherently leads to $\bar{X}$ being independent of $\hat{X}$ and $\bar{Y}$ being independent of $\hat{Y}$. Since the illuminated sample and dark sample are taken separately at two different points in time one can also reasonably assume the two samples to be independent of each other such that $\bar{X}$, $\hat{X}$, $\bar{Y}$, and $\hat{Y}$ are mutually independent; hence the photon induced mean signal and signal variance are represented by the independent variables

$$\bar{P} \sim \mathcal{N}\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right),$$
$$\hat{P} \sim \mathcal{GD}\left(\frac{n_1 - 1}{2}, \frac{n_2 - 1}{2}, \frac{n_1 - 1}{2\sigma_1^2}, \frac{n_2 - 1}{2\sigma_2^2}\right),$$

and the ratio $G = \bar{P}/\hat{P}$ is a centralized inverse-Fano random variate.

4.3 Principal-valued expectation, bias, and consistency of G

We now extend the notion of the principal-valued moment to multivariate functions of random variables. In the case of the centralized inverse-Fano variable, $PVE(G)$ can be found by first writing the density of G as

$$f_G(g) = \int_{\mathbb{R}} |\hat{p}|\, f_{\bar{P},\hat{P}}(g\hat{p}, \hat{p})\, \mathrm{d}\hat{p}. \tag{13}$$

Knowing that the non-zero density of $\hat{P}$ at the origin causes the expected value of G to diverge, a disk of radius δ centered at the origin is removed from the interval of integration in (13), which subsequently permits $PVE(G)$ to be written in the form

$$PVE(G) = \lim_{\delta \to 0} \int_{\mathbb{R}} g \int_{|\hat{p}|>\delta} |\hat{p}|\, f_{\bar{P},\hat{P}}(g\hat{p}, \hat{p})\, \mathrm{d}\hat{p}\mathrm{d}g.$$

Defining the transformation $T : \hat{p} = v,\ g = u/v$ one finds the Jacobian $|\det(\mathbf{J}(g,\hat{p}))| = |v|^{-1}$. Changing variables and simplifying subsequently yields

$$PVE(G) = E(\bar{P})\, PVE(\hat{P}^{-1}). \tag{14}$$

Then noting that $E(\bar{P}) = \mu_1 - \mu_2$ and that $\hat{P}^{-1}$ is an inverse gamma-difference random variable with $\alpha_i = (n_i - 1)/2$ and $\beta_i = \alpha_i/\sigma_i^2$ for $i = 1, 2$, the result of Theorem 3.3 can be used to produce an explicit expression for $PVE(G)$.

From the form in (14) one observes that the principal-valued expectation of G is not equal to the desired quantity $E(\bar{P})\,E(\hat{P})^{-1}$ as is to be expected. By generalizing the concept of expectation via the principal value we can quantify the principal-valued bias of the estimator G by $PVB(G) := PVE(G - g)$. With a bit of algebraic manipulation one finds

$$\begin{aligned} PVB(G) &= E(\bar{P})\big[PVE(\hat{P}^{-1}) - E(\hat{P})^{-1}\big] \\ &= E(\bar{P}) PVB(\hat{P}^{-1}). \end{aligned}$$

To characterize the asymptotic bias and consistency of G, we map all ordered pairs of sample sizes $(n_1, n_2) \in \mathbb{N} \times \mathbb{N}$ to the index $j \in \mathbb{N}$ via Cantor's pairing function $\pi(n_1, n_2) := \frac{1}{2}(n_1 + n_2)(n_1 + n_2 + 1) + n_2$ and then define the sequence of estimators $\{G_j\}_{j \geq 0}$ where $G_j = \bar{P}_j/\hat{P}_j$. Then, noting that $\bar{P}_j \xrightarrow{p} E(\bar{P})$ and $\hat{P}_j \xrightarrow{p} E(\hat{P})$, by Slutsky's theorem one finds for the probability limit

$$\operatorname*{plim}_{j \to \infty} G_j = E(\bar{P})\, E(\hat{P})^{-1} = g,$$

provided $E(\hat{P}) \neq 0$. This establishes that G is weakly consistent. Consequently, it is possible to control the bias of G at a given illumination level by varying the sample sizes n_1 and n_2. To illustrate this point, consider the expression for $PVE(G)$ in (14) and use the definition $g = E(\bar{P})E(\hat{P})^{-1}$ to write the relative bias of G as

$$\frac{PVE(G)}{g} = \frac{PVE(\hat{P}^{-1})}{E(\hat{P})^{-1}}. \tag{15}$$

We see that the relative bias of G w.r.t. g is just the relative bias of $\hat{P}^{-1}$ w.r.t. $E(\hat{P})^{-1}$. Furthermore, inspection of the r.h.s. of (15) and the definitions of α_i and β_i reveals that the relative bias of G is a function of the sample sizes n_i and populations variances σ_i^2, $i = 1, 2$. The zero illumination population variance σ_2^2 is a fixed quantity inherent to the pixel design while σ_1^2 is fixed for a fixed level of illumination; thus, the relative bias at a given illumination level is controlled exclusively by the sample sizes.

Table 1. Parameters used in simulation.

Parameter	Value(s)	Unit
g	2.1	e-/DN
μ_2	42.3	DN
σ_2	6.33	DN
μ_{e}-	(see figures 2–3)	e-
μ_1	$\mu_2 + \mu_{e}\text{-}/g$	DN
σ_1	$(\sigma_2^2 + \mu_{e}\text{-}/g^2)^{1/2}$	DN

To demonstrate how (15) can be used to control the relative experimental bias of g in a realistic scenario, the mean dark signal μ_2, dark noise σ_2, and conversion gain g were estimated from a pixel of a scientific grade charge-coupled device sensor as presented in Table 1. These estimates were treated as population parameters and subsequently used to generate the illuminated population parameters μ_1 and σ_1 based on the electron signal $\mu_{e}\text{-} = 400\,e\text{-}$ which represents the mean number of electrons generated by incident illumination on the pixel. This mean electron signal was chosen as to reflect an illumination level that puts the sensors output below the shot noise-limited regime. Equipped with all the required parameters, two normally distributed pseudo-random samples $\mathbf{X} = \{X_i\}_{i=1}^{2001}$ and $\mathbf{Y} = \{Y_i\}_{i=1}^{2001}$ where $X_i \sim \mathcal{N}(232.8, (11.4)^2)$ and $Y_i \sim \mathcal{N}(42.3, (6.33)^2)$ were generated using MATLAB software. Starting with the first $n_i = 21$ observations ($\alpha_i = 10$) in each sample, the variance of each population was estimated by $\hat{X}$ and $\hat{Y}$ and then $\alpha_1 = (n_1 - 1)/2$, $\alpha_2 = (n_2 - 1)/2$, $\beta_1 = \alpha_1/\hat{X}$, and $\beta_2 = \alpha_2/\hat{Y}$ were substituted into the r.h.s. of (15) to estimate the relative bias of G at the current sample sizes. This process was then iterated where for each iteration two more observations were added to the initial sample to re-evaluate the relative bias of G at the new sample sizes. Figure 2 shows the result of three runs of this experiment along with the theoretical bias (solid line) as a function of α_i. One can see that the estimated relative bias provides a reasonable approximation to the exact value especially for $\alpha_i > 100$. As such, an experimenter could follow a similar process whereby a shutter is placed in front of the sensor to switch between capturing dark and illuminated observations and then observe when the sample sizes become large enough to achieve a prescribed relative bias. Once the desired maximum relative bias is achieved the experiment would be terminated.

It is important to highlight the use of equal sample sizes in this demonstration. While a convenient first choice, equal sample sizes are unlikely to be optimal in the sense that there likely exists a pair of unequal sample sizes to achieve the same bias with a smaller total number of observations. The results and methodology presented here serve as a jumping off point for developing a procedure to iteratively monitor the relative bias while also controlling the relative size of each sample as to reach the desired bias with the minimum total number of observations.

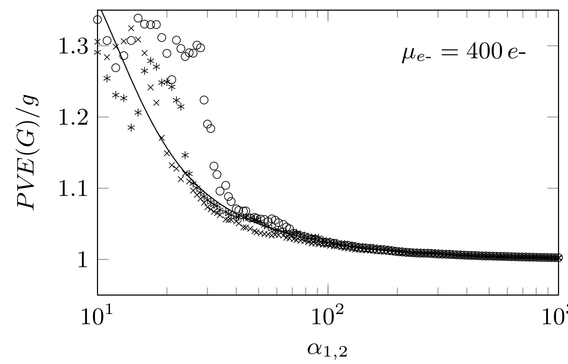

Figure 2. Relative bias $PVE(G)/g$ (solid) vs. $\alpha_1 = \alpha_2$ along with estimates for three Monte Carlo trials.

4.4 An inverse gamma-difference approximation to the CIF distribution

Consider the variances of the variables $\bar{P}$ and $\hat{P}$:

$$\begin{aligned} \mathrm{var}(\bar{P}) &= \sigma_1^2/n_1 + \sigma_2^2/n_2, \\ \mathrm{var}(\hat{P}) &= 2\sigma_1^4/(n_1 - 1) + 2\sigma_2^4/(n_2 - 1), \end{aligned}$$

where σ_2 is the pixel noise in the absence of illumination, $\sigma_1^2 = \sigma_2^2 + \sigma_\gamma^2$ is the total pixel noise variance under illumination, and σ_γ is the photon noise observed by the pixel. Each of these quantities is assumed to be in units of DN. It can be concluded that given sufficiently large σ_2 and σ_γ, the variance of $\hat{P}$ dominates that of $\bar{P}$. Under such circumstances it is then reasonable to treat $\bar{P}$ as degenerate w.r.t. $\hat{P}$, i.e. $\bar{P} \sim \delta(E(\bar{P}))$, which in turn leads to the approximation

$$G \approx E(\bar{P}) \times \hat{P}^{-1}. \tag{16}$$

Clearly this approximation is a scaled inverse gamma-difference variable which has the benefit of obviating the need for numerical integration to compute the approximate density of G. That said, another approximation is realized by normalizing each side of (16) by $PVE(G)$ yielding

$$\frac{G}{PVE(G)} \approx \frac{\hat{P}^{-1}}{PVE(\hat{P}^{-1})}, \tag{17}$$

provided $PVE(G) \neq 0$. By normalizing in this manner, all quantities relating to $\bar{P}$ have been removed; hence, the random variable $\hat{P}^{-1}/PVE(\hat{P}^{-1})$ has the additional benefit of not requiring the parameters μ_1 and μ_2. To get a better sense for the utility of this approximation, we again consider the parameters in Table 1. Using these parameters as well as the mean electron signals $\mu_{e}\text{-} = 1$, 2, and $10\,e$-, the distributions of $G/PVE(G)$ and $\hat{P}^{-1}/PVE(\hat{P}^{-1})$ were computed and plotted together as seen in Figure 3. As the illumination is increased (indicated by increasing μ_{e}-), the

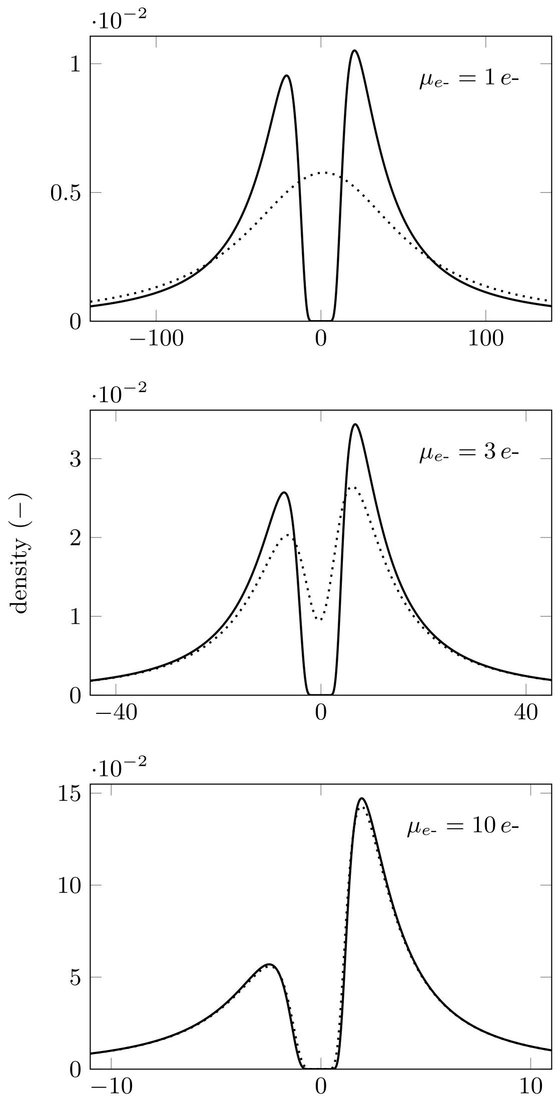

Figure 3. Density of $G/PVE(G)$ (dotted) vs. $\hat{P}^{-1}/PVE(\hat{P}^{-1})$ (solid) for the parameters presented in Table 1 and $\alpha_1 = \alpha_2 = 75$.

two distributions converge until they become nearly indistinguishable at a mean electron signal of $10\,e$-. The electron signals used here were chosen as to magnify the discrepancy between the two distributions and are much lower than that which would be used in a real experiment. As such, this example demonstrates how the distributions of $G/PVE(G)$ and $\hat{P}^{-1}/PVE(\hat{P}^{-1})$ can under appropriate conditions be assumed identical for the practical purpose of conversion gain measurement.

One reason for studying this approximation in more detail is that it leads to constructing relative confidence intervals for G. If the illumination is sufficiently high as to achieve the convergence seen in Figure 3 and the sample sizes are large then $g/G \approx \hat{P}/E(\hat{P})$ which is approximately normally distributed. Consequently, computing an interval estimate of $\hat{P}/E(\hat{P})$ and inverting the interval bounds should yield and approximate confidence interval for G/g so long as $\text{pr}(G \leq 0) \approx 0$. This relative interval would have practical use in design of experiment algorithms designed to estimate the sample sizes needed to measure the conversion gain to some desired relative uncertainty in a live experiment [10]. That said, more work has to be done to analyze how the magnitude of σ_2 affects the illumination level at which the convergence seen in Figure 3 takes places. It is clear that larger σ_2 will lead to convergence at lower illumination levels; however, rules of thumb for determining an appropriate illumination level for a given σ_2 are needed to make this more practical.

4.5 The low illumination problem

This paper introduced a theorem for computing principal-valued moments, the inverse gamma-difference distribution, its principal-valued first moment, and an application to electro-optical imaging sensor characterization via the photon transfer method. Being derived from the principals of normal sampling theory, the result contained herein are likely to be applicable to other fields of interest. That said, this work has a very specific higher goal in view.

Recall the application section where the discussion centered around the measurement of the conversion gain of sensors that exhibit a linear transfer function. For such devices, the gain is independent of signal level and can be measured at arbitrarily large illumination levels where undesirable characteristics of the gain measurement can be avoided. However, in practice, linearity tends to be an overly idealized model of the transfer function that is frequently violated by many different sensor designs and architectures. In fact, for the case of the increasingly commonplace complementary metal-oxide semiconductor (CMOS) active-pixel sensor (APS) technology, nonlinearity of the transfer function is an inherent consequence of the physics dictating the operation of the device.

In response to a need to characterize nonlinear sensors like the CMOS APS, Janesick [11] devised an extension to the photon transfer method, later dubbed the nonlinear compensation (NLC) technique [21], which makes a less restrictive assumption on the transfer function. Specifically, Janesick argued that the physics underlying many electro-optical sensors such as the CMOS APS permit a quasi-linear response at low levels of illumination; thus permitting the estimation of the conversion gain using the traditional formula. Once this initial measurement of the gain was performed, it is multiplied by subsequent measurements taken at several additional (higher) levels of illumination to effectively track a modified yet analogous gain estimator as a function of signal. While the justification for this approach is reasonable, properly measuring the gain at low illumination presents significant practical challenges that must be addressed.

First, at low illumination one must precisely control the sample sizes to ensure that the measurement of $\hat{P}$ remains positive and away from the origin. If this is not done properly, the gain measurement will be heavily biased and ill-behaved which are direct consequences of the nonlinearity of the transformation $\bar{P}/\hat{P}$ and its discontinuity at $\hat{P}=0$, respectively. To complicate matters more, the sample sizes required to maintain this balance also become very large at low illumination which motivates the need for a procedure to determine optimal sample sizes as discussed in section 4.3. Second, to make confidence intervals for G tractable we require that the uncertainty in the measurement of G be dominated by that of $\hat{P}$. While we have established some initial observations that will be useful to solve this problem, more work has to be done to determine rules of thumb for setting the illumination level based on the magnitude of signal independent noise present in the sensor. In short, these two problems constitute what we call the *low illumination problem*. The results derived in this paper open the door to the analysis needed to solve this important practical problem and subsequently develop a more robust and generalized approach to the photon transfer method.

APPENDIX A. ALTERNATIVE PROOF FOR THEOREM 3.2

Proof. We begin with the random variable X with density function f_X that is continuously differentiable and $|\frac{\mathrm{d}}{\mathrm{d}x} f_X(x)| < M$ holds in a neighborhood of the origin for $M \in \mathbb{R}$. The moments of X partitioned about the origin are defined by

$$E(X_0^{\epsilon-1}) = E(X^{\epsilon-1}\mathbb{1}_{Y\le 0}) + E(X^{\epsilon-1}\mathbb{1}_{Y>0}).$$

Writing each expected value using the integral definition leads to

$$\lim_{\epsilon\to 0} E(X_0^{\epsilon-1}) = \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{0} + \int_{0}^{\infty}\right) x^{\varepsilon-1} f_X(x)\,\mathrm{d}x.$$

Next, we define the branch cut $z^\varepsilon = |z|^\varepsilon e^{\mathrm{i}\varepsilon \arg z}$ where $-\pi/2 < \arg z < 3\pi/2$ and then write

$$\lim_{\epsilon\to 0} E(X_0^{\epsilon-1}) = \lim_{\varepsilon\to 0}\left(\int_{0}^{\infty} x^{\varepsilon-1} f_X(x)\,\mathrm{d}x - e^{\mathrm{i}\pi\varepsilon}\int_{-\infty}^{0} |x|^{\varepsilon-1} f_X(x)\,\mathrm{d}x\right).$$

With a bit of algebraic manipulation, the integrand of the second term can be written in terms of the nascent delta function $\eta_\varepsilon(x) = \varepsilon|x|^{\varepsilon-1}$.

$$\lim_{\epsilon\to 0} E(X_0^{\epsilon-1}) = \lim_{\varepsilon\to 0}\int_{\mathbb{R}} |x|^\varepsilon \frac{f_X(x)}{x}\,\mathrm{d}x - \pi\mathrm{i}\lim_{\varepsilon\to 0}\left(\frac{e^{\mathrm{i}\pi\varepsilon}-1}{\mathrm{i}\pi\varepsilon}\right)\int_{-\infty}^{0} \eta_\varepsilon(x) f_X(x)\,\mathrm{d}x$$

In the limit as $\varepsilon \to 0$, η_ε converges to the Dirac delta function in the sense that

$$\lim_{\varepsilon\to 0}\int_{-\infty}^{0} \eta_\varepsilon(x) f_X(x)\,\mathrm{d}x = f_X(0).$$

Thus we have

$$\lim_{\epsilon\to 0} E(X_0^{\epsilon-1}) = \lim_{\varepsilon\to 0}\int_{\mathbb{R}} |x|^\varepsilon \frac{f_X(x)}{x}\,\mathrm{d}x - f_X(0)\pi\mathrm{i}. \tag{18}$$

For the remaining term we first partition the interval of integration by[2]

$$\int_{\mathbb{R}} |x|^\varepsilon \frac{f_X(x)}{x}\,\mathrm{d}x = \int_{|x|>\delta} |x|^\varepsilon \frac{f_X(x)}{x}\,\mathrm{d}x + f_X(0)\int_{-\delta}^{\delta} \frac{|x|^\varepsilon}{x}\,\mathrm{d}x + \int_{-\delta}^{\delta} |x|^\varepsilon \frac{f_X(x)-f_X(0)}{x}\,\mathrm{d}x. \tag{19}$$

Starting with the first term on the right hand side of (19) it is straightforward to show that the limit can be evaluated directly according to the dominated convergence theorem. For the second term we see that it evaluates to zero when $\varepsilon > 0$ due to the integrand being odd and the interval of integration being symmetric about the origin. Therefore, this term adds no contribution in the limit. Lastly, since f_X is continuously differentiable on $\mathbb{R}$, the singularity of the integrand at the origin when $\varepsilon = 0$ is removable. Again calling on the dominated convergence theorem the limit of this term is evaluated directly. Altogether, the limit yields

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}} |x|^\varepsilon \frac{f_X(x)}{x}\,\mathrm{d}x = \int_{|x|>\delta} \frac{f_X(x)}{x}\,\mathrm{d}x + \int_{0}^{\delta} \frac{f_X(x)-f_X(-x)}{x}\,\mathrm{d}x,$$

which we recognize as the Cauchy principal value. Substituting this result back into (18) yields the expression

$$\lim_{\epsilon\to 0} E(X_0^{\epsilon-1}) = \mathrm{PV}\int_{\mathbb{R}} \frac{f_X(x)}{x}\,\mathrm{d}x - f_X(0)\pi\mathrm{i}.$$

Upon, adding $f_X(0)\pi\mathrm{i}$ to both sides, replacing $E(X_0^{\epsilon-1})$ by its definition, and identifying the remaining integral as $PVE(X^{-1})$, the desired result is found. □

APPENDIX B. ALTERNATIVE DERIVATION OF THE FIRST NEGATIVE MOMENT OF THE NORMAL DISTRIBUTION

The principal-valued first negative moment of the normal density function was first introduced in a paper by Quenouille where it was derived by the means of a differential

[2] The author thanks Daniel Fischer for suggesting this approach on `math.stackexchange.com` [22].

equation [15]. Here we provide an abbreviated alternative derivation for this expression through application of Theorem 3.2.

Proof. We begin by defining $X \sim \mathcal{N}(\mu, \sigma^2)$ and then partitioning the integral representation for the moments $E(X^{\varepsilon-1})$ about the origin. From [23, Eq. 3.462.1] the partitioned moments are found to be

$$E(X^{\varepsilon-1}\mathbb{1}_{X\leq 0}) = -e^{\mathrm{i}\pi\varepsilon} f_X(0)\exp\left(\tfrac{\mu^2}{4\sigma^2}\right)\Gamma(\varepsilon)D_{-\varepsilon}\left(\tfrac{\mu}{\sigma}\right),$$
$$E(X^{\varepsilon-1}\mathbb{1}_{X>0}) = f_X(0)\exp\left(\tfrac{\mu^2}{4\sigma^2}\right)\Gamma(\varepsilon)D_{-\varepsilon}\left(-\tfrac{\mu}{\sigma}\right),$$

where $D_\nu(z)$ is the parabolic cylinder function. By definition, the sum of these two expressions yields the moments of the normal density function which is continuous in ε.

$$E(X_0^{\varepsilon-1}) = f_X(0)\exp\left(\tfrac{\mu^2}{4\sigma^2}\right)\times \Gamma(\varepsilon+1)\frac{D_{-\varepsilon}\left(-\frac{\mu}{\sigma}\right) - e^{\mathrm{i}\pi\varepsilon}D_{-\varepsilon}\left(\frac{\mu}{\sigma}\right)}{\varepsilon}$$

As $\varepsilon \to 0$, the leading gamma term tends to unity while the ratio is indeterminant. Applying L'Hôpitals rule the limit becomes

$$\lim_{\varepsilon\to 0} E(X_0^{\varepsilon-1}) = f_X(0)\exp\left(\tfrac{\mu^2}{4\sigma^2}\right)\times \lim_{\varepsilon\to 0}\tfrac{\partial}{\partial\varepsilon}\Big(D_{-\varepsilon}\left(-\tfrac{\mu}{\sigma}\right) - e^{\mathrm{i}\pi\varepsilon}D_{-\varepsilon}\left(\tfrac{\mu}{\sigma}\right)\Big).$$

According to [24, Eq. 07.41.20.0005.01]

$$\tfrac{\partial}{\partial\nu}D_\nu(z)\Big|_{\nu=0} = \frac{1}{2}e^{-z^2/4}\Bigg(-{}_2F_2\left(\begin{matrix}1,1\\ \frac{3}{2},2\end{matrix};\tfrac{z^2}{2}\right)z^2 + \pi\,\mathrm{erfi}\left(\tfrac{z}{\sqrt{2}}\right) - \log 2 - \gamma\Bigg),$$

where $\mathrm{erfi}(z)$ is the imaginary error function and $\gamma = 0.57721\ldots$ is the Euler–Mascheroni constant. Making use of this result as well as the relation $\mathcal{D}(z) = -\frac{1}{2}\mathrm{i}\sqrt{\pi}e^{-z^2}\,\mathrm{erf}(\mathrm{i}z)$, where $\mathcal{D}(z)$ is the Dawson integral, we arrive at the final expression for the limit.

$$\lim_{\varepsilon\to 0} E(X_0^{\varepsilon-1}) = \tfrac{\sqrt{2}}{\sigma}\mathcal{D}\left(\tfrac{\mu}{\sqrt{2}\sigma}\right) - f_X(0)\pi\mathrm{i}$$

Dropping the imaginary component then yields the solution derived by Quenouille for $PVE(X^{-1})$. ☐

ACKNOWLEDGEMENTS

The author would like to thank Dr. Mostafa Aminzadeh and Mr. Raymond Rogers for their thoughtful and constructive critiques of the work contained in this paper. The author also thanks Dr. Chien-Yu Peng and Dr. Bernhard Klar for the fruitful discussions on negative moments and gamma-difference distributions, respectively.

Received 25 June 2018

REFERENCES

[1] Klar, B., A note on gamma difference distributions. *Journal of Statistical Computation and Simulation*, 85(18):3708–3715, Feb 2015. MR3401104

[2] Press, S. J., On the sample covariance from a bivariate normal distribution. *Annals of the Institute of Statistical Mathematics*, 19(1):355, Dec 1967. MR0219160

[3] Kotz, S., Kozubowski, T., and Podgorski, K., *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Progress in Mathematics Series. Birkhäuser Boston, 2001. MR1935481

[4] Holm, H. and Alouini, M. S., Sum and difference of two squared correlated nakagami variates in connection with the mckay distribution. *IEEE Transactions on Communications*, 52(8):1367–1376, Aug 2004.

[5] Augustyniak, M. and Doray, L. G., Inference for a leptokurtic symmetric family of distributions represented by the difference of two gamma variates. *Journal of Statistical Computation and Simulation*, 82(11):1621–1634, 2012. MR2984565

[6] Mathai, A.M., On noncentral generalized Laplacianness of quadratic forms in normal variables. *Journal of Multivariate Analysis*, 45(2):239 – 246, 1993. MR1221920

[7] Krishna, E. and Kanichukattu, J., Marshall-Olkin generalized asymmetric Laplace distributions and processes. *Statistica*, 71(4):453–467, 2011.

[8] Omura, J. and Kailath, T., *Some useful probability distributions*. Stanford University, 1965.

[9] Simon, M. K., *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers, Scientists and Mathematicians*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[10] Hendrickson, A. J., Centralized inverse-Fano distribution for controlling conversion gain measurement accuracy of detector elements. *J. Opt. Soc. Am. A*, 34(8):1411–1423, Aug 2017.

[11] Janesick, J. R., *Photon transfer: DN → λ*. SPIE, 2007.

[12] Beecken, B. P. and Fossum, E. R., Determination of the conversion gain and the accuracy of its measurement for detector elements and arrays. *Appl. Opt.*, 35(19):3471–3477, Jul 1996.

[13] NIST, *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.

[14] Piegorsch, W. W. and Casella, G., The existence of the first negative moment. *The American Statistician*, 39(1):60–62, 1985. MR0784670

[15] Quenouille, M. H., Notes on bias in estimation. *Biometrika*, 43(3/4):353–360, Dec 1956. MR0081040

[16] Peng, C. Y., The first negative moment in the sense of the Cauchy principal value. *Statistics & Probability Letters*, 78(13):1765–1774, September 2008. MR2453915

[17] Peng, C. Y., The first negative moment of skew-t and generalized Students t-distributions in the principal value sense. *Journal of Applied Mathematics*, 2013:1–12, 2013. MR3096023

[18] Shapiro, J. M., Domains of attraction for reciprocals of powers of random variables. *SIAM Journal on Applied Mathematics*, 29(4):734–739, 1975. MR0383506

[19] Kang, H. and An, C., Differentiation formulas of some hypergeometric functions with respect to all parameters. *Applied Mathematics and Computation*, 258(Supplement C):454–464, 2015. MR3323082

[20] Dereniak, E. L. and Crowe, D. G., *Optical radiation detectors*. J. Wiley, 1984.

[21] Bohndiek, S. E., Blue, A., Clark, A. T., Prydderch, M. L., Turchetta, R., Royle, G. J., and Speller, R. D., Comparison of methods for estimating the conversion gain of cmos active pixel sensors. *IEEE Sensors Journal*, 8(10):1734–1744, Oct 2008.

[22] Fischer, D. (math.stackexchange.com/users/83702/daniel fischer). An approach to evaluating a Cauchy principal value that yields unexpected extra imaginary term. Mathematics Stack Ex-

change. URL: https://math.stackexchange.com/q/2745257 (version: 2018-04-20).

[23] Gradshteyn, I. S. and Ryzhik, I. M., *Table of Integrals, Series, and Products*. Academic Press, 8 edition, 2014. MR0669666

[24] Formula 07.41.20.0005.01. The wolfram functions site. Visited on 07/01/18.

Aaron Hendrickson
Atlantic Test Ranges, Code 5.2.4.4.1
23013 Cedar Point Road, Patuxent River
Maryland 20670-1183, U.S.A
E-mail address: ajh4184@gmail.com