Consistent Optical and Electrical Noise Figure

Reinhold Noe

***Abstract*—The optimum noise figure of an electrical amplifier is $F_e = 1$ and the optimum traditional noise figure of an optical amplifier is $F_{pnf} = 2$. This irresolvable conflict is due to the fact that F_e measurement requires electrical powers, proportional to squared amplitudes (voltages), while F_{pnf} measurement requires squares and variances of photocurrents, proportional to 4th powers of amplitudes (fields). In line with this an electrical receiver can receive I&Q parts of an electric carrier while a direct-detection photoreceiver can detect only power and not phase of an optical carrier. Optical amplifiers cause Gaussian field noise. Photodetection causes shot noise. Spontaneous-spontaneous beat noise in direct detection is taken into account by negative binomial or chi-squared photoelectron distributions, without errors of a Gaussian approximation. Coherent receivers linearly sense the Gaussian field noise. The sensitivity of an ideal coherent I&Q receiver is not degraded if it gets an ideal optical preamplifier, while the corresponding $F_{pnf} = 2$ suggests degradation. Coherent I&Q or heterodyne receivers have electrical output powers proportional to squared amplitudes (fields). This way one has the same metric in electrical and optical domain. One gets an optical noise figure $F_{o,IQ}$. For large amplifier gain it is $F_{pnf}/2$. In an ideal amplifier, $F_{o,IQ} = 1$. For true optical homodyne receivers and for direct detection receivers with Gaussian approximation it can be converted into F_{pnf} and vice versa. Phase-sensitive amplifiers are also covered. With F_e and the I&Q optical noise figure $F_{o,IQ}$ a consistent unified noise figure is derived, valid and usable in electrical and optical domain.**

***Index Terms*—Noise figure, optical amplifiers, optical fiber communication.**

I. Introduction

The noise behavior of a linear electrical twoport device at frequency f and reference temperature T can be described by its gain G and electrical noise figure F_e [1], [2]. Measured powers are proportional to squares of amplitudes (voltages). It is near at hand to use the same metric for optical devices. But the traditional optical noise figure given by EDFA pioneer E. Desurvire [3], which may be written as $F_{pnf} = 1/G + 2\tilde{\mu}$ at high powers, has a conflicting definition: Squares and variances of measured photocurrents are proportional to squares of optical powers, i.e., 4th powers of amplitudes (optical fields). I noticed this discrepancy and have therefore used in my lectures since 1996 [4] a quantity which I now call $\tilde{\mu}$ and which is the expectation value of input-referred noise photons per mode. It has the same cascading properties as the electrical excess noise figure $F_e - 1$. At that time I did not question the validity of the conflicting definition [3], all the more given that it also fulfills Friis' casdading formula.

Noise figure pioneer H. A. Haus independently questioned Desurvire's noise figure, called it F_{pnf} (photon number fluctuations) and published two noise figures [5], [6]. Upon invitation [6] he defined combined usage in electrical and optical domains. The two noise figures are both claimed to become identical to F_e in the electrical domain but are otherwise different. This suggests that at least one of them is inappropriate.

One is F_{fas}[1] based on fluctuations of amplitude squares, eqn. (18) of [6]. In the optical domain $kT << hf$ it approaches F_{pnf} at high optical powers. This suggests that F_{fas} is based on amplitudes taken to the 4th power, not on amplitude squares. In the electrical domain the spontaneous emission factor n_{sp} ($\equiv \theta$ in [5], [6]) is not given and should scale inversely with frequency. If F_{fas} approaches F_e for $kT >> hf$ then it carries the same conceptual conflict as F_{pnf} vs. F_e. If it doesn't then it is in value conflict with F_e.

The other is F_{ASE} [5], [6] based on amplified spontaneous emission, also here with n_{sp} unknown at low frequencies. So I published [7], [8] about an $F = 1 + \tilde{\mu} \equiv F_{ASE}$ in the optical domain. Photoelectron distributions are calculated from the input-referred number $\tilde{\mu}$ of noise photons per mode and other parameters [7], also for D(Q)PSK [8].[2] The good thing about $\tilde{\mu}$ and F_{ASE} is that measured powers are proportional to the squares of amplitudes (fields), like for F_e. This requires a receiver which is linear in amplitude (voltage, optical field). In the optical domain that is a coherent receiver. But F_{ASE} does not describe SNR degradation in a coherent receiver.

Desurvire's F_{pnf} has proven extremely useful since decades. But firstly it was derived for direct-detection receivers. These have meanwhile been replaced by coherent I&Q receivers in most cases in which optical amplifier noise matters. Secondly, F_{pnf} has prevented an optical noise figure definition that is consistent with the definition of F_e. H. A. Haus' F_{fas} and his (and my) F_{ASE} have problems, too, and are not established. All three approach 2 for an ideal optical amplifier with $n_{sp} = 1$ and $G \to \infty$.

A coherent optical I&Q receiver is linear from optical input amplitudes to electrical output amplitudes in both quadratures, just like an electrical amplifier is linear from input to output

Manuscript received 3 June 2022; revised 22 July 2022 and 1 September 2022; accepted 6 September 2022. Date of publication 10 October 2022; date of current version 1 January 2023.

The author is with the EIM-E, Paderborn University, 33098 Paderborn, Germany and also with Novoptel GmbH, 33100 Paderborn, Germany (e-mail: noe@upb.de).

This article has supplementary material provided by the author and color versions of one or more figures available at https://doi.org/10.1109/JLT.2022.3212936.

Digital Object Identifier 10.1109/JLT.2022.3212936

[1]Note that the middle result and the last result in (13) of [5] and the final definition (18) of [6] evaluated in the optical domain $kT << hf$ are three different, unequal versions of F_{fas}. The middle result in (13) of [5], which was abandoned in favor of the later (18) of [6], equals $F_{o,IQ}$, which is advocated in this paper.

[2]The technical content of this paper shall also enter a possible next edition of [8].

in both quadratures. The sensitivity of an ideal coherent I&Q receiver is not degraded if it gets an ideal optical preamplifier. In striking contrast, a noise figure equal to 2 (as predicted by $F_{pnf} = F_{fas}$ and F_{ASE}) suggests degradation.

The above conflicts must be resolved for consistent definition and usage of electrical and optical noise figure.

The electrical noise figure is reviewed in Section II. The existing optical noise figures are reviewed in Section III. Photoelectron statistics are reviewed in Section IV. and in the Appendix. Noise figures in coherent optical receivers are derived in Section V. How the foregoing changes for phase-sensitive amplifiers is discussed in Section VI. All noise figures are compared in Section VII. The noise figure $F_{o,IQ}$ for a coherent optical I&Q or heterodyne receiver conceptually matches the electrical F_e. In Section VIII. the definitions are combined into a consistent noise figure F_{IQ} that is valid in electrical and optical domains. Maybe it should be called *F*.

II. Review of Electrical Noise Figure

We define

$$G \;=\; \text{(available) power gain,} \tag{1}$$

$$P_{s,in} = \text{input signal power,} \tag{2}$$

$$P_{s,out} = GP_{s,in} = \text{ output signal power.} \tag{3}$$

The gain *G* applies at the operation frequency *f*. It can for instance be determined by measurement of $P_{s,in}$ and $P_{s,out}$. A source with known noise, such as a noisy source that can be put at two temperatures, may avoid the need of providing an extra signal. For simplicity let us assume: Source, device (amplifier or attenuator) and detector have the same temperature *T*. Source, device and detector be impedance-matched. In this case thermal noise powers do not depend on impedances.

We want to determine powers in a one-sided, physical

$$B = 1/\tau = \text{bandwidth.} \tag{4}$$

With a power detector, usually incorporating a preamplifier with large gain, one can measure the

$$P_{n,in} = kTB = \text{ input noise power.} \tag{5}$$

It is thermal noise. In spite of detector noise, *kTB* can be determined or inferred, by heating or cooling the source to different temperatures.

In later context it is expedient to use a filter with bandwidth *B* and a rectangular envelope of the sinusoidal impulse response. If τ is the length of this envelope then $B = 1/\tau$ is the bandwidth (4).

Given that $P_{n,in} = kT/\tau$ is the noise input power, *kT* is the noise energy collected in time τ. We can call this the thermal noise energy in one mode. Compared to a signal carrier, the noise can be in phase and in quadrature. So, *kT*/2 is the thermal energy per mode and quadrature. A mode contains 2 quadratures.

Number of modes times quadratures per mode equals the number of degrees-of-freedom. In the baseband or at DC there is only one quadrature. In agreement with the foregoing the voltage variance at a capacitance in parallel with a thermally noisy conductance is *kT*/*C* (= kT/C noise) [9] and the mean stored energy is *kT*/2.

Furthermore we need

$$P_{n,out} = \text{ output noise power,} \tag{6}$$

$$SNR_{in} = \frac{P_{s,in}}{P_{n,in}/2} = \text{ input signal to noise ratio,} \tag{7}$$

$$SNR_{out} = \frac{P_{s,out}}{P_{n,out}/2} = \text{ output signal to noise ratio.} \tag{8}$$

Here the noise powers are divided by 2 because only half the noise is in phase with the signal. This is not mandatory (and is not possible in thermal detectors). But it will facilitate discussion of single-quadrature cases found in optics. With or without this division one arrives at the standard definition

$$F = \frac{SNR_{in}}{SNR_{out}} = \frac{P_{n,out}}{GP_{n,in}} = \text{ noise figure,} \tag{9}$$

$$F_{ex} = F - 1 = \text{ excess noise figure,} \tag{10}$$

$$T_{ex} = (F-1)\,T = \text{ excess noise temperature.} \tag{11}$$

An electrical amplifier has the electrical noise figure

$$F_e = \frac{P_{n,out}}{GP_{n,in}} = \frac{P_{n,out}}{GkTB} > 1. \tag{12}$$

The added thermal noise output power $GkT_{ex}B$ is independent of generator temperature. If one reduces generator temperature *T*, total output noise power $P_{n,out} = Gk(T + T_{ex})B$ is reduced while $F_e = 1 + T_{ex}/T$ is increased. On the other hand, cooling of the amplifier generally reduces its T_{ex} (for instance because its internal resistances generate less thermal noise) and reduces both $P_{n,out}$ and F_e.

For an electrical attenuator it holds

$$P_{n,out} = kTB, \tag{13}$$

$$F_e = 1/G. \tag{14}$$

When the attenuator has the same temperature *T* as the generator it behaves like the generator and delivers the same $P_{n,out} = kTB$. If the attenuator has the temperature T_{att} then $P_{n,out} = kTGB + kT_{att}(1-G)B$. The generator noise is attenuated according to $G < 1$, and for $T_{att} = T$ we get the usual $P_{n,out} = kTB$. Cooling of generator or attenuator reduces $P_{n,out}$. Cooling of generator increases, cooling of attenuator reduces $F_e = 1 + (T_{att}/T)(1/G - 1)$.

Electrical power measurent in the carrier or radio frequency domain yields total power in both quadratures. Using two mixers with local oscillator signals having $\pi/2$ mutual phase difference one can just as well downconvert this to the baseband. Addition of the squares of the resulting I&Q signals gives total power in the two quadratures. If the receiver is a heterodyne receiver it needs an image rejection filter at the input. Otherwise doubled noise power would be measured (for a given input noise power spectral density).

III. Review of Optical Noise Figures

A purely amplifying optical medium has an optical power change $dP/dt = aP$ where a is relative stimulated emission coefficient per time unit. For pure attenuation we get $dP/dt = -bP$ where b is the relative attenuation coefficient per time unit. For concurrent amplification and attenuation power gain equals

$$G = e^{(a-b)t} = e^{(a-b)z/v_g}. \quad (15)$$

Here t is time needed to pass the amplifier, z is its length and v_g is group velocity.

At the amplifier output one finds added noise in this mode with

$$P_n\tau = \mu hf = G\tilde{\mu}hf = n_{sp}\,(G-1)\,hf = \text{ energy per mode,} \quad (16)$$

$$n_{sp} = \frac{a}{a-b} = \text{ spontaneous emission factor.} \quad (17)$$

If we want to consider the optical amplifier as noiseless then μ photons per mode must be added at the amplifier output, or $\tilde{\mu}$ at the input. Tilde $^\sim$ means input-referred.

For an amplifier with equal gain and loss ($a = b$, $G = 1$, $1/n_{sp} = 0$) we get

$$\begin{aligned}\mu = n_{sp}\,(G-1) &= \lim_{a\to b}\frac{a}{a-b}\left(e^{(a-b)t}-1\right)\\ &= at = bt = az/v_g = bz/v_g = n_{sp}\,(1-1/G) = \tilde{\mu}. \quad (18)\end{aligned}$$

E. Desurvire [3] has defined the signal-to-noise ratio as

$$SNR_{pnf} = \langle n\rangle^2\Big/\sigma_n^2 \quad (19)$$

The suffix means photon number fluctuations and has been given by H. A. Haus [5], [6].

A sufficiently attenuated optical signal has a Poisson distribution of the photoelectron number detected during a time τ. The variance of a Poisson distribution equals the mean value. This means

$$SNR_{pnf,in} = \langle n\rangle^2\Big/\langle n\rangle. \quad (20)$$

At the amplifier output, the mean signal equals $G\langle n\rangle$. For large $\langle n\rangle$ the variance is roughly $G\langle n\rangle + 2n_{sp}(G-1)G\langle n\rangle$. Under this approximation one obtains output SNR and Desurvire's and Haus' noise figure

$$SNR_{pnf,out} = \frac{G^2\langle n\rangle^2}{G\,\langle n\rangle + 2n_{sp}\,(G-1)\,G\,\langle n\rangle}, \quad (21)$$

$$\begin{aligned}F_{pnf} &= \frac{SNR_{pnf,in}}{SNR_{pnf,out}}\\ &= \frac{\langle n\rangle^2}{\langle n\rangle}\frac{G\,\langle n\rangle + 2n_{sp}\,(G-1)\,G\,\langle n\rangle}{G^2\langle n\rangle^2}\\ &= \frac{1+2n_{sp}\,(G-1)}{G} = 1/G + 2\tilde{\mu}. \quad (22)\end{aligned}$$

For finite $\langle n\rangle$, the resulting exact F_{pnf} (not given in this paper) varies with $\langle n\rangle$, i.e., optical power, and also depending on the mode number, i.e., polarizations (1 or usually 2) and ratio $B_o/(2B_e)$ of optical to doubled electrical bandwidth. Anyway, spontaneous-spontaneous beat noise is approximately taken into account by the exact, power-dependent expression for F_{pnf}. This was advantageous when people wanted to avoid the more complicated exact BER calculation (see next Section and Appendix) and when coherent receivers were not used. But today it is useless or even disadvantageous because the standard usage of optical amplifier chains is coherent I&Q transmission. No-one would accept an F_e which, like the exact F_{pnf}, depended on signal power and detector structure.

Providing a broadband direct detection receiver for measurement of F_{pnf} is costly. Optical power measurement is much cheaper.

H. Haus [6] has defined a noise figure, intended to evaluate fluctuations of amplitude squares, which in its final version, and evaluated in the optical domain, equals

$$F_{fas} = 1/G + 2\tilde{\mu}, \quad (23)$$

identical with F_{pnf} (at high powers).

He has also defined a noise figure based on amplified spontaneous emissions [6], which, when evaluated in the optical domain, can be written as

$$F_{ASE} = 1 + \tilde{\mu}. \quad (24)$$

A linear optical receiver, i.e., a coherent receiver, is needed to measure electrical powers proportional to noise powers. But while its (electrical) signal power is proportional to G its (electrical) noise power is not proportional to GF_{ASE}. Namely, shot noise generated in a coherent receiver is G times smaller when referred to the amplifier input. This means F_{ASE} does not describe SNR degradation.

IV. Short Review of Photoelectron Statistics

Bit error ratio (BER) calculation of optical receivers is laid out in [8] and – for direct-detection ASK receivers only – in [7]. This closely follows the fundamental work [10], [11]. Mathematical help is in [12]. The most important parts of this are compiled in the Appendix of this manuscript. In all cases, knowledge of $\tilde{\mu}$, the input-referred number of noise photons per mode, is needed. A merit of F_{ASE} is that its excess is readily $\tilde{\mu} = F_{ASE} - 1$.

Assume an optical signal with constant power. During a time interval τ it has the energy $\mu_0 hf$. The signal is detected in a photodiode which for simplicity has a quantum efficiency of 1. The number of photoelectrons is Poisson-distributed with the (generally non-integer) expectation value μ_0. If the power is switched between different values we get a weighted addition of Poisson distributions with different expectation values. We define an intensity x equal to a possibly variable or random value of μ_0. It has the probability density function (PDF) $p_x(x)$. The probability $P(n)$ of detecting n photoelectrons equals [11]

$$P\,(n) = \int_0^\infty p_x\,(x)\,e^{-x}\frac{x^n}{n!}dx \text{ (Poisson transform)}. \quad (25)$$

For constant power, $p_x(x) = \delta(x - \mu_0)$ is a Dirac and the Poisson transform delivers a Poisson distribution with expectation value μ_0. The signal has a constant intensity, and the photodetection or the Poisson transform adds shot noise.

The master equation of photon statistics is a differential equation for the time-dependent evolution of $P(n)$ while light is subject to spontaneous and stimulated emission and absorption. For the moment generating function of $P(n)$ a differential equation can be derived. Its solution allows calculating $P(n)$ after the signal has passed a device [11]. If the device is a strong attenuator then the output $P(n)$ is a Poisson distribution. If it is an amplifier with non-zero or zero input power then $P(n)$ is a non-central or central negative binomial distribution, respectively, with $2N$ degrees-of-freedom or N modes. The Poisson transform of a chi-squared variable is a negative binomial distribution. This means the photoelectron distribution behind the amplifier belongs to a chi-squared intensity (that is subject to photodetection / shot noise generation). A chi-squared variable x with $2N$ degrees-of-freedom ($= \chi^2_{2N}$ distribution) is the sum of $2N$ squares of statistically independent Gaussian variables with equal variances.

Now let us look at the intensity x processed by a photoreceiver behind an optical amplifier. For simplicity, we define optical fields $\mathbf{E}$ to have the unit $\sqrt{W}$ such that the squared field magnitude equals the power (not probability)

$$P = |\mathbf{E}|^2. \tag{26}$$

The amplifier input signal has a normalized field

$$\mathbf{E}_{in} = \sqrt{\tilde{\mu}_0/M}\underline{\mathbf{e}}_1 e^{j\omega t}\sqrt{hf/\tau_o}. \tag{27}$$

The normalized Jones vector is $\underline{\mathbf{e}}_1$, the angular optical frequency is $\omega = 2\pi f$. The photon number arriving in a detection time interval τ has the expectation value $\tilde{\mu}_0$. Quantity $M = \tau/\tau_o \approx B_o/(2B_e)$ is the number of statistically independent photocurrent samples during τ. There is an optical filter with bandwidth B_o. In this model the optical impulse response should have a rectangular envelope of length $\tau_o = 1/B_o$ and the optical receiver should simply add M photocurrent samples with temporal neighbor spacings τ_o, to cover the detection interval $\tau \approx 1/(2B_e)$ where B_e is the resulting electrical bandwidth.[3]

The input signal is passed through an optical amplifier with gain G and then through the optical filter. The output signal has a field $\mathbf{E}_{out} = \sqrt{G}\mathbf{E}_{in}$. Let us assume the optical amplifier adds field noise $((u_1 + ju_2)\underline{\mathbf{e}}_1 + (u_3 + ju_4)\underline{\mathbf{e}}_2)e^{j\omega t}$. The second normalized Jones vector $\underline{\mathbf{e}}_2$ is orthogonal to $\underline{\mathbf{e}}_1$. The $u_i = \sqrt{G}\tilde{u}_i$ ($i = 1 \ldots 4$) are zero-mean in-phase and quadrature noises in the $p = 2$ polarizations supported by amplifier and receiver. Referred to the input of the amplifier, which may now be assumed to be noiseless, the total field is

$$\tilde{\mathbf{E}} = \left(\left(\sqrt{\tilde{\mu}_0/M} + \tilde{u}_1 + j\tilde{u}_2\right)\underline{\mathbf{e}}_1 + (\tilde{u}_3 + j\tilde{u}_4)\underline{\mathbf{e}}_2\right)e^{j\omega t} \times \sqrt{hf/\tau_o}. \tag{28}$$

[3]This is an approximation. Strictly speaking the electrical bandwidth of the receiver which just sums photocurrent samples is infinite. But the decision variable x contains energy collected in time interval τ.

Let $R = e/(hf)$ be the (ideal) photodiode responsivity. Photocurrent is $I = RP$. A normalized photocurrent sample is

$$I(t) = R|\mathbf{E}|^2 = G\left(\left(\sqrt{\tilde{\mu}_0/M} + \tilde{u}_1\right)^2 + \tilde{u}_2^2 + \tilde{u}_3^2 + \tilde{u}_4^2\right) \times (e/\tau_o). \tag{29}$$

The receiver sums up M independent photocurrent samples, and multiplies the result by τ_o/e, in order to form a dimensionless decision variable $x = G\tilde{x}$. Referred to the amplifier input (i.e., divided by G) it is

$$\tilde{x} = \frac{\tau_o}{Ge}\sum_{i=1}^{M} I(t + i\tau_o). \tag{30}$$

But we have to consider the photodetector. The photoelectron distribution is given by the Poisson transform of $x = G\tilde{x}$.

We know already that it is negative binomial distributed. This means that x is a variable with chi-square PDF. The only possible interpretation is that indeed statistically independent Gaussian field noises $u_i = \sqrt{G}\tilde{u}_i$ with equal variances are added by the amplifier!

The optical signal with constant power comes into the amplifier as a wave. It is important to understand that optical or quantum noise can be subdivided into two issues, amplifier noise and shot noise:

- The amplifier adds Gaussian field noise in phase and in quadrature of each mode.
- Photodetection adds shot noise. For large photocurrents, shot noise can be neglected.

Indeed, for large expectation values (caused by large gain or large input power), a non-central or a central negative binomial or a Poisson distribution of x approaches a non-central or a central chi-squared or a Dirac distribution, respectively.

In order to avoid scaling of distributions it is useful to evaluate the input-referred $\tilde{x}$. It has a noncentral negative binomial (or Laguerre) distribution with expectation value $\tilde{\mu}_0$ of the signal, expectation value $\tilde{\mu}$ of noise photons per mode and $2N = 2pM$ degrees-of-freedom or N modes. All $\tilde{u}_i$ have the variance $\sigma^2 = \tilde{\mu}/2$. If there is no optical amplifier (and $G = 1$) then the noncentral negative binomial distribution degenerates into a Poisson distribution. If there is no signal, $\tilde{\mu}_0 = 0$, it degenerates into a central negative binomial distribution.

For transmitted 1 or 0 there are different values of $\tilde{\mu}_0$. The resulting two distributions allow determining the optimum threshold and the bit error ratio (BER).

If there is also Gaussian electrical noise then the distributions of $\tilde{x}$ need to be convolved with a Gaussian distribution with a variance equal to that of of the expected charge number generated by the circuit in τ. The detection time interval τ is chosen a bit shorter than the symbol duration, in order to minimize intersymbol interference.

If the direct detection receiver is for DPSK or DQPSK then differences of photocurrents are formed at the outputs of a delay-line interferometer. Sums and differences of time-delayed independent noise samples can be transformed into new independent noise samples. BER can again be calculated.

So, it suffices to know modulation format, receiver structure, expectation value $\tilde{\mu}_0$ of signal photoelectrons that would be detected at the optical amplifier input in time interval τ, optical filter impulse response length τ_o, gain G and $\tilde{\mu}$, the input-referred number of noise photons number per mode of the amplifier.

No knowledge of spontaneous-spontaneous beat noise is needed and the result is more exact than if spontaneous-spontaneous beat noise had been approximated by a Gaussian distribution! The key for exact receiver performance calculation is knowledge of $\tilde{\mu}$ (and N).

If a coherent receiver is not balanced then the above needs to be considered also there. But usually it is balanced and there will be no direct detection (or spontaneous-spontaneous beat noise) effect.

For simplicity let us further assume that the local oscillator power approaches infinity, $P_{LO} \to \infty$. Then, in one quadrature of one polarization, there is a Gaussian decision variable with expectation value $\sqrt{\tilde{\mu}_0}$ and standard deviation $\sqrt{\tilde{\mu}/2}$. Details will be discussed in the next Section. Unless P_{LO} is small, shot noise can be neglected and noise is purely Gaussian, given that the field noise of optical amplifiers is Gaussian. For synchronous detection, BER calculation is trivial (erfc). For asynchronous detection the same calculus can be applied to the electrical signals as we have applied to the optical fields in direct detection receivers. So, also for coherent receivers, $\tilde{\mu}$ is needed and is decisive.

Quantum effects with largest impact are

- Gaussian field noise of optical amplifier in front of any receiver, and
- shot noise of strong signal in direct detection receiver without optical amplifier,
- shot noise of local oscillator in coherent receiver without optical amplifier.

V. Noise Figures in Coherent Optical Receivers

We want to extend F_e to all frequencies including the optical domain. A necessary condition is that an optical noise figure F_o can be obtained from measurements which are physically equivalent to those needed for F_e. Electrical power is proportional to the square of an amplitude (voltage or current). This means optical power must be measured to determine F_o. Optical power is proportional to the square of an amplitude (electric or magnetic field).

Photocurrent is proportional to optical power. The electrical power of a photocurrent is proportional to its square, hence proportional to the 4th power of an amplitude. Clearly, powers of photocurrents cannot be used to determine F_o.

In the electrical domain, amplitudes (voltages) are measured directly, and devices (amplifiers) are linear. One could do the same in the optical domain. But direct optical amplitude (field) measurement is not possible. A linear translation from optical to electrical amplitudes is hence needed. The electrical output amplitude of a (linear) coherent optical receiver is proportional to an optical field amplitude. So, for determining F_o, signal and noise powers can be measured proportional to the electrical power at the output of a coherent receiver.

An electrical signal can be upconverted into the optical domain by an I&Q modulator. An optical signal can be downconverted into an electrical signal by a coherent receiver.

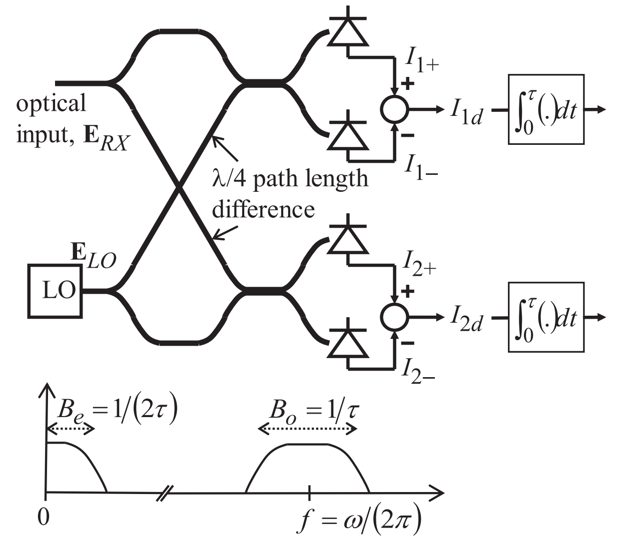

Fig. 1. Coherent I&Q receiver with polarization matching (top), spectrum (bottom).

In order to work like an electrical device or amplifier the coherent receiver must be a receiver for both quadratures. This can be a synchronous heterodyne receiver with image rejection filter at the input. For convenience we take instead an equivalent I&Q receiver, composed of a 3 dB power splitter and two homodyne receivers with local oscillator signals having $\pi/2$ mutual phase difference (Fig. 1 top).

We again define $P = |\mathbf{E}|^2$. The polarizations be identical, with normalized Jones vector $\mathbf{e}_1$. We assume electric fields

$$\begin{aligned} \mathbf{E}_{RX} &= \sqrt{G}\left(\sqrt{P_s} + (v_1 + jv_2)\sqrt{P_n/2}\right)\mathbf{e}_1 e^{j\omega t} \\ \mathbf{E}_{LO} &= j\sqrt{P_{LO}}\mathbf{e}_1 e^{j\omega t} \end{aligned} \tag{31}$$

with signal power P_s, equivalent input noise power P_n, local oscillator power P_{LO} and independent zero-mean Gaussian in-phase and quadrature noises v_1, v_2 having unity variance $\sigma^2 = 1$. The detected photocurrents and their differences and sums are

$$\begin{aligned} I_{1\pm} &= R|\pm\mathbf{E}_{RX}/2 + \mathbf{E}_{LO}/2|^2 \\ &= \frac{R}{4}\left(G\left(P_s + 2v_1\sqrt{P_sP_n/2} + \left(v_1^2 + v_2^2\right)P_n/2\right)\right. \\ &\qquad \left. \pm 2\left(\sqrt{P_s} + v_1\sqrt{P_n/2}\right)\sqrt{GP_{LO}} + P_{LO}\right) \\ I_{2\pm} &= R|\pm\mathbf{E}_s/2 + j\mathbf{E}_{LO}/2|^2 \\ &= \frac{R}{4}\left(G\left(P_s + 2v_1\sqrt{P_sP_n/2} + \left(v_1^2 + v_2^2\right)P_n/2\right)\right. \\ &\qquad \left. \pm\, 2v_2\sqrt{P_n/2}\sqrt{GP_{LO}} + P_{LO}\right), \end{aligned} \tag{32}$$

$$\begin{aligned} I_{1d} &= I_{1+} - I_{1-} = R\left(\sqrt{P_s} + v_1\sqrt{P_n/2}\right)\sqrt{GP_{LO}} \\ I_{2d} &= I_{2+} - I_{2-} = Rv_2\sqrt{P_n/2}\sqrt{GP_{LO}} \\ I_{1s} &= I_{1+} + I_{1-} = R\Big(G\Big(P_s + 2v_1\sqrt{P_sP_n/2} \\ &\qquad + \left(v_1^2 + v_2^2\right)P_n/2\Big) + P_{LO}\Big)/2 \end{aligned}$$

$$I_{2s} = I_{2+} + I_{2-} = R\Big(G\Big(P_s + 2v_1\sqrt{P_sP_n/2} + \left(v_1^2+v_2^2\right)P_n/2\Big) + P_{LO}\Big)/2, \tag{33}$$

where $R = e/(hf)$ is the (ideal) photodiode responsivity. One-sided shot noise power spectral densities in I_{1d}, I_{2d} are $2eI_{1s}$ and $2eI_{2s}$. In an ideal receiver it holds $P_{LO} \to \infty$. Hence we may neglect the term $G\left(P_s + 2v_1\sqrt{P_sP_n/2} + \left(v_1^2+v_2^2\right)P_n/2\right)$ in $2eI_{1s}$, $2eI_{2s}$, also thermal receiver noise. Let us assume the coherent receivers integrates the currents over a time τ. This corresponds to an electrical bandwidth $B_e = 1/(2\tau)$ (Fig. 1 bottom). The optical bandwidth in which the I&Q receiver detects noise is $B_o = 2B_e = 1/\tau$. The noise generated by the optical amplifier has a one-sided power spectral density $P_n/B_o = \tilde{\mu}hf$. This is the input-referred noise energy per mode. For SNR calculation I_{2d}, I_{2s} are not needed. At the output we obtain the SNR

$$\begin{aligned} SNR_{o,IQ,out} &= \frac{\overline{I_{1d}}^2}{\sigma_{I_{1d}}^2 + \sigma_{I_{1s}}^2} \\ &= \frac{R^2GP_sP_{LO}}{R^2P_{LO}\tilde{\mu}GhfB_o/2 + eRP_{LO}B_e} \\ &= \frac{RP_s\tau}{R\tilde{\mu}hf/2 + e/(2G)} = \frac{P_s\tau}{(\tilde{\mu}+1/G)\,hf/2}. \end{aligned} \tag{34}$$

Under synchronous detection (namely QPSK, QAM), BER is given by erfc functions. In the I&Q receiver asynchronous detection (ASK, DPSK, DQPSK) is also possible. BER calculation is the same as for the corresponding direct detection receivers (comprising 0, 1 or 2 interferometers) as outlined in Section IV. (with details given for instance in [8] and the Appendix of this paper).

In a true homodyne receiver there are no power splitters and there is only receiver branch 1, for the in-phase signal. Hence P_s, P_n and P_{LO} arrive in this receiver (branch) twice as strong as in Fig. 1. We get

$$\begin{aligned} SNR_{o,I,out} &= \frac{4R^2GP_sP_{LO}}{4R^2P_{LO}\tilde{\mu}GhfB_o/2 + 2eRP_{LO}B_e} \\ &= \frac{2P_s\tau}{(2\tilde{\mu}+1/G)\,hf/2}. \end{aligned} \tag{35}$$

The quantum noise energy per mode and quadrature is *hf*/2. Assume the (mean) signal energy $P_s\tau = \mu_0 hf$ corresponds to an average of μ_0 detected photons and the added noise power is zero, $P_n\tau = \tilde{\mu}hf = 0$, because there is no optical amplifier. All the same the resulting $SNR_{o,IQ} = 2\mu_0$, $SNR_{o,I} = 4\mu_0$ are finite. Shot noise, here (mainly) that of the local oscillator, limits SNR. Also in direct detection SNR is limited, see (20).

In the coherent receiver P_{LO} dominates. Its photodetection causes shot noise, even if input powers are zero, $P_s = P_n = 0$. It is hence appropriate to model an internal shot noise source in the coherent receiver. Shot noise is *hf* per mode when referred to the receiver input.

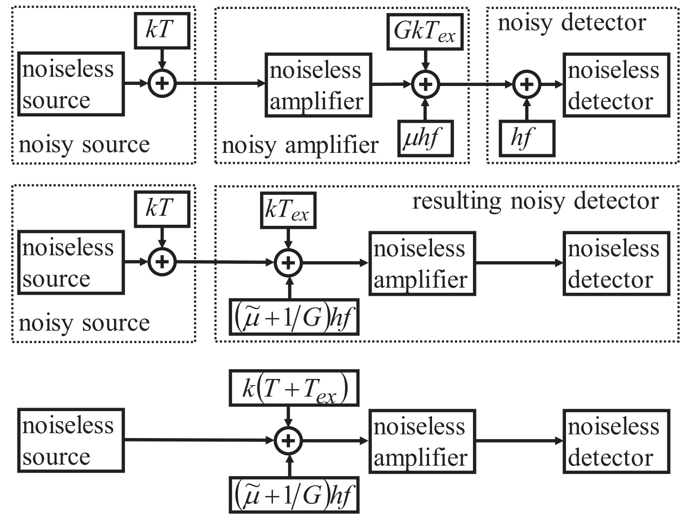

Fig. 2. Source, amplifier and detector (linear I&Q, electrical or coherent optical), noisy or noiseless with equivalent added noise energies per mode. Individual devices (top) and equivalent interpretations (middle, bottom).

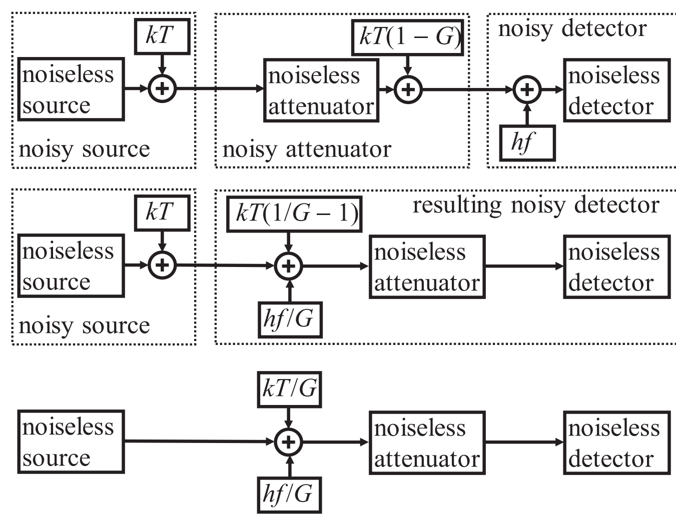

Fig. 3. Like Fig. 2, but for attenuator instead of amplifier.

When an amplifier or other device is inserted in front of the receiver it becomes part of a new combined detector, as will be seen in Figs. 2 and 3.

To evaluate (34) without optical amplifier we set $G = 1$, $n_{sp} = 0$ (or finite), $\tilde{\mu} = n_{sp}(1-1/G) = 0$ and obtain

$$SNR_{o,IQ,in} = \frac{P_{s,in}\tau}{hf/2}, \tag{36}$$

$$\begin{aligned} F_{o,IQ} &= \frac{SNR_{o,IQ,in}}{SNR_{o,IQ,out}} = \tilde{\mu} + 1/G \\ &= \frac{n_{sp}(G-1)+1}{G} \\ &= 1 + (n_{sp}-1)(1-1/G). \end{aligned} \tag{37}$$

Let us check the last expression. In an amplifier, $G \geq 1$, $1 - 1/G \geq 0$, $n_{sp} - 1 \geq 0$. This means $F_{o,IQ} \geq 1$.

We can convert F_{pnf} (in the limit of high powers) into the optical I&Q noise figure $F_{o,IQ}$ and vice versa,

$$\begin{aligned} F_{o,IQ} &= (F_{pnf} - 1/G)/2 + 1/G \\ F_{pnf} &= 2\,(F_{o,IQ} - 1/G) + 1/G. \end{aligned} \tag{38}$$

In the usual case $G >> 1$,

$$F_{o,IQ} \approx F_{pnf}/2 \qquad F_{o,IQ,\mathrm{dB}} \approx F_{pnf,\mathrm{dB}} - 3\mathrm{dB} \tag{39}$$

where $F_{\ldots,\mathrm{dB}}$ is a noise figure expressed in dB.

Conversion from/to F_{ASE} is likewise possible,

$$\begin{aligned} F_{o,IQ} &= F_{ASE} - 1 + 1/G \\ F_{ASE} &= F_{o,IQ} - 1/G + 1. \end{aligned} \tag{40}$$

For the true homodyne receiver we obtain

$$SNR_{o,I,in} = \frac{2P_s\tau}{hf/2}, \tag{41}$$

$$F_{o,I} = \frac{SNR_{o,I,in}}{SNR_{o,I,out}} = 2\tilde{\mu} + 1/G = \frac{2n_{sp}\,(G-1)+1}{G}. \tag{42}$$

The optical true homodyne noise figure $F_{o,I}$ is identical with F_{pnf} (for high powers) even though the receivers are quite different. All the same, the result $F_{o,I} = F_{pnf}$ (for high power) does not surprise too much because the true homodyne receiver and the intensity-modulation-direct-detection (IM-DD) receiver with high power both don't evaluate quadrature noise. Compared to the I&Q receiver the true homodyne receiver has no power splitter at the input. This is why its shot noise, referred to the input, is only half as large. As a consequence, $SNR_{o,I,out}$ and $F_{o,I}$ are affected twice as much by $\tilde{\mu}$ than $SNR_{o,IQ,out}$ and $F_{o,IQ}$.

The foregoing can for instance be verified with BER measurements. As mentioned, optical amplifiers generate Gaussian noise of field and of coherent receiver output signals.

Coherent receivers cut off nasty noise in extra modes, namely a ratio $B_o/(2B_e) > 1$ and the second polarization $\mathbf{e}_2$. A dual-polarization coherent receiver processes each of the two signal polarizations like a single-polarization coherent receiver processes a single signal polarization.

$SNR_{o,IQ}$, $SNR_{o,I}$, $F_{o,IQ}$, $F_{o,I}$ are defined for coherent optical receivers. But coherent receivers are not experimentally required: One can measure G and $p\mu B_o$ simply with an optical power meter and an optical filter having a known bandwidth B_o. If the amplifier is polarization-insensitive there are $p = 2$ polarizations. If the amplifier is polarizing then we have $p = 1$ polarization. Of course, time-gated noise measurement during short dark periods of an applied optical signal (or another method) is required to maintain the wanted correct G (which, due to saturation, is lower than the idle G). From the obtained μ and G one can calculate $F_{o,IQ}$, $F_{o,I}$.

The coherent $SNR_{o,IQ,in}$, $SNR_{o,IQ,out}$ are halved if noise from both quadratures is taken. This does not change their quotient $F_{o,IQ}$. All the same it seems more appropriate to take noise into account only in phase with the signal:

- The best possible signal processing involves synchronous detection, which cuts off quadrature noise.
- Insertion of a phase-sensitive amplifier would lead to an unfair comparison if noise of both quadratures were evaluated.

VI. Phase-Sensitive Amplifiers

Let us check how the foregoing changes when optical phase-sensitive amplifiers (PSA) [13], [14], [15] are used, namely degenerate parametric optical amplifiers. For simplicity let the gain be sufficiently high, $G >> 1$. In (34), (35) the term $\tilde{\mu}G = \mu$, multiplied by $hf/2$, is only the in-phase noise. It must be substituted by the (ideally) half as high in-phase noise of the PSA. In spite of physical differences, let us consider $\tilde{\mu}$ simply as a quantity which ideally can be as low as 1. If we do so and want to describe the PSA then we must replace $\tilde{\mu}$ by $\tilde{\mu}/2$ in (34), (35) and also in (37), (42).

As a result, the insertion of an ideal optical PSA before a true homodyne receiver will give it a noise figure $\tilde{\mu} + 1/G$. This is equal to the expression $F_{o,IQ}$ of the I&Q receiver. It can approach the value 1, identical to the traditional noise figure value $F_{pnf} = 1$ of an ideal optical PSA. One sees that the high sensitivity of the true homodyne receiver is not deteriorated by the ideal PSA.

If the ideal PSA is inserted before an ideal optical I&Q receiver it will have a noise figure $\tilde{\mu}/2 + 1/G$. In the limit it can be as low as 1/2. This means: The I&Q receiver sensitivity, which is 2 times worse than that of the true homodyne receiver, can be improved, times 2, to the true homodyne receiver sensitivity at the input of an inserted optical PSA.

For high gain $G >> 1$ of the optical PSA the noise of the subsequent receiver can be neglected, no matter whether it is direct detection, homodyne or I&Q. In all cases one gets the homodyne sensitivity at the input of the optical PSA.

This fits into the foregoing considerations, there is no contradiction. Logically, an ideal single or dual quadrature amplifier has an ideal noise figure $F_{o,I} = F_{pnf} = F_{fas}$ or $F_{o,IQ}$ of 1 before a single or dual quadrature receiver, respectively.

Note that there is a difference between optical and electrical PSA, which has nothing to do with noise figure definition. An optical PSA is quantum-based. This is why at high G the in-phase noise is halved, and sensitivity for the in-phase signal of an I&Q receiver is ideally doubled. In electrical PSA [16], governed by thermal noise, one defines the noise squeezing factor. The best possible value equals 1/2. In that case there is only in-phase noise and no quadrature noise. In-phase noise has the same gain as the signal. When looking at only the in-phase components, SNR stays unchanged (is not doubled by the electrical PSA) and ideal noise figure equals 1.

VII. Comparison of Noise Figures

We start by deriving a condition (43) for noise figures. In order to consider the device as noiseless, we say a noise μ (of unspecified kind) is added at the device output or a noise $\tilde{\mu} = \mu/G$ is added at its input. For a cascade of devices 1, 2 it holds $\mu = \mu_1 G_2 + \mu_2$, $\tilde{\mu} = \tilde{\mu}_1 + \tilde{\mu}_2/G_1$ and $G = G_1 G_2$. Let

TABLE I
COMPARISON OF NOISE FIGURES

Type of noise figure F	F of ideal ampl.	F of atten., $G < 1$	M of ampl.	Input-referred energy per mode, kT_{ex} or $\tilde{\mu}hf$
F_e	1	$1/G$	≥ 0	$kT(F-1)$
$F_{o,IQ} = n_{sp}(1-1/G)+1/G$	1	$1/G$	$n_{sp}-1 \geq 0$	$hf(F-1/G)$
$F_{pnf} = F_{fas} = F_{o,I}$ $= 2n_{sp}(1-1/G)+1/G$	2	$1/G$	$2n_{sp}-1$ ≥ 1	$hf(F-1/G)/2$
$F_{ASE} = 1+n_{sp}(1-1/G)$	2	1	$n_{sp} \geq 1$	$hf(F-1)$

us assume there exists a noise figure defined by

$$F = A + (1 - A)/G + B\tilde{\mu}. \tag{43}$$

A, *B* are constants. Term *A* stands for noise generated in the source, namely thermal noise. Term $(1 - A)/G$ stands for noise generated upon detection, namely shot noise. *B* brings device noise into the game. One easily shows that the device cascade fulfills

$$F = F_1 + (F_2 - 1)/G_1. \tag{44}$$

By complete induction one shows that definition (43) fulfills Friis' cascading formula for arbitrary device cascades,

$$F - 1 = \sum_{i=1}^{n} \frac{F_i - 1}{\prod_{k=1}^{i-1} G_k}. \tag{45}$$

As can be seen, excess noise figures are added after having been divided by the total prior gain.

Equation (43) gives some freedom in noise figure design. For F_e it holds $A = 1$ and for instance $B = 1$, $\tilde{\mu} = T_{ex}/T$ where T_{ex} is an excess noise temperature. All the above noise figures F_e, $F_{pnf} = F_{fas} = F_{o,I}$, F_{ASE}, $F_{o,IQ}$ fulfill (43)!

One can cascade many optical amplifiers and attenuating fiber links. Transmission with a certain BER is possible if the signal power needed before the first amplifier is not too large. This is found using the noise figure of the whole cascade.

Ascending noise measure

$$M = \frac{F - 1}{1 - 1/G} \tag{46}$$

indicates the order in which amplifiers with different noise figures and gains should be cascaded for lowest total noise.

We compare the various noise figures (Table I). In all cases, input-referred noise energy per mode can be calculated from the noise figure. For the ideal optical amplifier we assume not only $n_{sp} = 1$ but also $G \to \infty$. Based on noise measure *M*, all optical noise figures yield the same cascading rule for noise minimization: Order amplifiers according to ascending n_{sp}. For an ideal amplifier, $M = 0$ in the case of F_e and $F_{o,IQ}$ but $M = 1$ in the case of $F_{pnf} = F_{fas} = F_{o,I}$ and F_{ASE}.

For a pure attenuator, F_{ASE} equals 1 whereas the other noise figures equal $1/G$. More importantly, calculation of $F_{o,IQ}$, $F_{o,I}$ has shown that the differing F_{ASE} is not an SNR degradation factor. Reason is that fundamental quantum noise (shot noise) is added upon photodetection, not by the source.

If one wanted to match $F_{pnf} = F_{fas} = F_{o,I}$ with F_e then, for increasing *f*, squares of amplitudes (these squares correspond to electrical RF powers) would need to gradually become 4th powers of (field) amplitudes (these 4th powers correspond to squares of photocurrents), or heterodyne receivers would need to gradually become homodyne receivers. This seems impossible.

Obviously $F_{o,IQ}$ is the only optical noise figure which matches F_e and avoids contradictions. The reason for this is that both are defined for the same kind of system (linear I&Q) with the same metrics (powers proportional to squared amplitudes).

For an ideal optical amplifier it holds $F_{o,IQ} = 1$. This is against traditional physical practice $F_{pnf} = 2$. But it is natural for communication engineers:

- An ideal optical preamplifier does not deteriorate the sensitivity of a system with ideal optical I&Q receiver.
- Ideal electrical amplifiers have the same low noise figure value $F_e = 1$.
- Just like electrical amplifiers, coherent optical I&Q receivers are linear in amplitudes and transparent for both quadratures. True optical homodyne receivers pass only one quadrature (unsuitable). Optical direct-detection receivers are not linear in (field) amplitudes (unsuitable) and, when ASK or DPSK, pass only one quadrature (unsuitable).
- Today, coherent optical I&Q receivers are standard in most cases in which optical noise figure matters, namely long-haul amplified fiber communication.
- $F_{o,IQ}$ of an amplifier or of a whole cascade of amplifiers and fibers directly tells how much worse I&Q receiver sensitivity is at the amplifier or cascade input. No manipulation (38) of F_{pnf} is needed.

VIII. CONSISTENT UNIFIED NOISE FIGURE

Given that F_e defines powers proportional to squares of amplitudes and F_{pnf} defines powers proportional to the squared squares of amplitudes there is no way to reconcile F_{pnf} with F_e.

One cannot say F_{pnf} is only for quantum detectors such as photodiodes and F_e is only for electrical detectors because the noise figure must not depend on the choice of detector. In particular it might one day be possible to construct electrical and quantum detectors for the same *f* (in the low THz range?) and this would oppose presumably unequal F_{pnf} and F_e for the same usage of the same device.

Electrical and optical signals are both electromagnetic signals and should be treated equally. The conceptual match of F_e and $F_{o,IQ}$ allows deriving a consistent unified noise figure, usable in electrical, thermal and optical domains.

Fig. 2 top depicts source, amplifier and detector, noisy or noiseless with equivalent added noises. The detector is linear in amplitude for I&Q components. If it is optical then it is a coherent I&Q receiver, comprising LO, couplers, photodiodes like in Fig. 1 and subsequent electrical amplifiers. Fig. 3 is like Fig. 2, but for an attenuator instead of an amplifier. The source generates thermal noise *kT*. Of course the detector also generates thermal noise. But that can be calibrated away after connecting the source directly to the detector and putting it to two different temperatures *T*. The detector adds shot noise *hf*.

The amplifier in Fig. 2 works at frequency *f*. In the electrical domain it generates thermal excess noise, GkT_{ex} at the output or kT_{ex} at the input. In the optical domain it adds field noise. This is $\mu hf = n_{sp}(G-1)hf$ at the output. All these noises are Gaussian (unless P_{LO} is small). Their powers can be added. In the thermal domain, both noises need to be considered.

Amplifier noise may be considered like a signal. If desired, it may be pushed over into the detector where it forms a total quantum noise $(\mu+1)hf$ (optical amplifer noise and shot noise). For $\mu >> 1$ the shot noise can be neglected. Total quantum noise $(\mu+1)hf$ can also conveniently be pushed to the amplifier input, of course divided by *G* because it is then amplified by *G*. Total quantum noise referred to the amplifier input is $(\tilde{\mu}+1/G)hf$ with $\tilde{\mu}=\mu/G$. Thermal noise of the amplifier can also been pushed to its input. This way we may consider the noisy amplifier and the noisy detector together as a new, resulting noisy detector, see Fig. 2 middle. If also the two thermal noise sources are combined then we arrive at Fig. 2 bottom which is recognized in the following equations.

Optical and electrical gains *G* are identical because they manifest at the same frequency *f*. Total thermal noise in bandwidth B_o at the amplifier output is $GF_e kTB_o$. Half of this is in phase with the signal. In the coherent I&Q receiver it appears multiplied with $R^2 P_{LO}$, like the amplified signal power GP_s. In (34) we add this thermal noise power σ_e^2,

$$\begin{aligned} SNR_{IQ,out} &= \frac{\overline{I_{1d}}^2}{\sigma_e^2+\sigma_{I_{1d}}^2+\sigma_{I_{1s}}^2} \\ &= \frac{R^2 GP_s P_{LO}}{R^2 P_{LO} GF_e kTB_o/2 + R^2 P_{LO}\tilde{\mu} GhfB_o/2 + eRP_{LO}B_e}. \\ &= \frac{GP_s}{GF_e kTB_o/2 + \tilde{\mu} GhfB_o/2 + hfB_e} \\ &= \frac{P_s\tau}{F_e kT/2 + F_{o,IQ}hf/2} = \frac{P_s\tau}{k(T+T_{ex})/2 + (\tilde{\mu}+1/G)hf/2} \end{aligned} \tag{47}$$

The 3rd line of (47), where $R^2 P_{LO}$ has been canceled, is the quotient of signal and noise powers in a possible I&Q electrical receiver (instead of I&Q receiver with quantum detectors).

For thermal detectors the same expression holds, except that noise powers will be twice as large because quadrature noise will not be suppressed. The same noise power doubling will also apply in the following $SNR_{IQ,in}$ (48). Hence F_{IQ} (49) will stay unchanged in thermal detectors.

Remember that noise energies *hf*/2, *kT*/2 are per quadrature whereas *hf*, *kT* in Figs. 2, 3 are per mode.

Without amplifier it holds $G=1, T_{ex}=0, n_{sp}=0$ (or finite), $\tilde{\mu}=0$. We get

$$SNR_{IQ,in} = \frac{P_s\tau}{kT/2+hf/2}, \tag{48}$$

$$\begin{aligned} F_{IQ} &= \frac{SNR_{IQ,in}}{SNR_{IQ,out}} = \frac{F_e kT + F_{o,IQ}hf}{kT+hf} \\ &= \frac{k(T+T_{ex}) + (\tilde{\mu}+1/G)hf}{kT+hf} \qquad \left(A = \frac{kT}{kT+hf}\right). \\ &= A + (1-A)/G + (AT_{ex}/T + (1-A)\tilde{\mu}) \end{aligned} \tag{49}$$

In case of an attenuator it holds $G<1$, $T_{ex}=T(1/G-1)$, $n_{sp}=0$, $\tilde{\mu}=0$, $F_e=F_{o,IQ}=F_{IQ}=1/G$.

As required, the unified I&Q noise figure F_{IQ} fulfills (43) and Friis' cascading formula. At low *f*, F_{IQ} becomes F_e. At high *f* it becomes $F_{o,IQ}$. Among the optical noise figure definitions unified with F_e, only F_{IQ} yields the same noise figure ($=1$) for ideal amplifiers in electrical and optical domains. The others, F_{fas}, F_{ASE} in [6] and F_{pnf} if unified with F_e, fail to fulfill this.

To determine F_{IQ} one must measure powers. At low *f* this needs an electrical measurement (preamplifier, squarer). At high *f* this is possible in a coherent optical receiver. And at all *f* one can just as well measure power directly, at least in principle and after preamplification, but including quadrature noise. Power can be measured as dissipated heat. In the optical domain it is more convenient to use an optical filter and an optical power meter, also with quadrature noise, than an I&Q or heterodyne receiver. So, power detection is always possible and detector choice is a matter of convenience, not a definition.

IX. Conclusion

It has been shown that all prior optical and unified noise figures lead into contradictions. Reasons are nonlinear metrics (of $F_{fas}=F_{pnf}$) which differ from linear metrics applied for the measurement of the electrical noise figure F_e or the not taking into account (in F_{ASE}) that shot noise generated in a coherent receiver is *G* times smaller when referred to the amplifier input. In the optical domain they all are misleading because an ideal amplifier has an $F_{fas}=F_{pnf}=2=F_{ASE}$ although it does not degrade the SNR of an ideal coherent optical I&Q receiver. The contradictions are resolved by an optical I&Q noise figure $F_{o,IQ}\geq 1$ which is defined with the same metrics as F_e: two quadratures, linear in amplitude, and powers are proportional to squares of amplitudes. For large gains $G>>1$, $F_{o,IQ}$ is simply $F_{pnf}/2$, i.e., 3 dB less if expressed in dB.

For true homodyne receivers an $F_{o,I}=F_{pnf}\geq 2$ is found. Phase-sensitive amplifiers can improve this to 1. Coherent optical I&Q systems with phase-sensitive amplifiers are made to operate like (more sensitive) homodyne systems.

In amplified fiber links, direct-detection receivers, for which an optimum noise figure $F_{pnf}=2$ made some sense, have been largely replaced by coherent I&Q receivers, which need $F_{o,IQ}$.

Combination of $F_{o,IQ}$ with F_e yields a unified I&Q noise figure F_{IQ}. It fulfills Friis' cascading formula. It does not depend on the detector type (electrical, thermal, optical). Maybe $F_{o,IQ}$ should be called F_o. Maybe F_{IQ} should be called *F*.

Appendix
BER Calculation in Optical Direct Detection Receivers

This derivation is from [8]. All random variables (RV) except fields be ≥ 0. For small time increments dt the probability $P(n,t+dt)$ to detect *n* photons in a medium at time $t+dt$

depends on the probabilities to detect $n-1, n$, or $n+1$ photons at time t, and on the conditional transition probabilities from one of these numbers to n,

$$P(n,t+dt) = P(n|n)\,P(n,t) + P(n|n-1) \times P(n-1,t) + P(n|n+1)\,P(n+1,t). \tag{50}$$

For $dt \to 0$ the transition probabilities are

$$P(n|n-1) = ((n-1)a+c)\,dt \qquad P(n|n+1) = (n+1)\,b\,dt$$
$$P(n|n) = 1 - P(n-1|n) - P(n+1|n) = 1-(n(a+b)+c)\,dt \tag{51}$$

with a = stimulated emission rate per photon, b = absorption rate per photon, c = spontaneous emission rate. This yields the

$$\begin{aligned} dP(n,t)/dt &= (P(n,t+dt) - P(n,t))/dt \\ &= -(n(a+b)+c)\,P(n,t) \\ &\quad + ((n-1)a+c)\,P(n-1,t) \\ &\quad + (n+1)\,bP(n+1,t) \end{aligned} \tag{52}$$

master equation of photon statistics.

If a particular photon statistic is to persist it must fulfill (52) but statistical parameters such as expectation value $\langle n\rangle$ may be time-variable. E.g., if there is only attenuation ($b>0$, $a=c=0$), a Poisson distribution

$$P(n) = e^{-\mu_0}\frac{\mu_0^n}{n!} \text{ with } \langle n\rangle = \mu_0(t) = \mu_0(0)\,e^{-bt} \tag{53}$$

results. It is conserved under pure attenuation.

The moment generating function (MGF) is

$$M_n\left(e^{-s}\right) = \left\langle e^{-sn}\right\rangle = \sum_{n=-\infty}^{\infty} P(n)\,e^{-sn} \text{ for discrete } n, \tag{54}$$

$$M_x\left(e^{-s}\right) = \left\langle e^{-sx}\right\rangle = \int_{-\infty}^{\infty} p_x(x)\,e^{-sx}dx \text{ for continuous } x. \tag{55}$$

The lower boundary can be set 0 for our nonnegative RVs. Inverse Laplace or, using $e^{-s}=z^{-1}$, inverse z transform allows obtaining probabilities $P(n)$ or probability density function $p_x(x)$. See Table II for important examples. Adding statistically independent RVs requires convolution of the corresponding PDFs, or multiplication of the corresponding MGFs. The MGF allows obtaining the moments,

$$\left\langle x^k\right\rangle = (-1)^k \left.\frac{d^k M\left(e^{-s}\right)}{(ds)^k}\right|_{s=0} \quad (\text{same for } \left\langle n^k\right\rangle). \tag{56}$$

$P(n,t)$ is time-variable while a signal passes an optical amplifier. We determine

TABLE II
SOME RV DISTRIBUTIONS WITH MGFS, MEAN VALUES AND VARIANCES

Discrete	$P(n)$ ($n\ge 0$)	$M_n\left(e^{-s}\right)$	$\langle n\rangle$	σ_n^2
Poisson	$e^{-\mu_0}\frac{\mu_0^n}{n!}$	$e^{-\mu_0\left(1-e^{-s}\right)}$	μ_0	μ_0
Central negative binomial	$\frac{\binom{n+N-1}{n}\mu^n}{(1+\mu)^{n+N}}$	$\frac{1}{\left(1+\mu\left(1-e^{-s}\right)\right)^N}$	$N\mu$	$N\mu(\mu+1)$
Non-central negative binomial, Laguerre	$\frac{\mu^n e^{-\frac{\mu_0}{1+\mu}}}{(1+\mu)^{n+N}} \cdot L_n^{N-1}\left(\frac{-\mu_0}{\mu(1+\mu)}\right)$	$\frac{e^{\frac{-\mu_0\left(1-e^{-s}\right)}{1+\mu\left(1-e^{-s}\right)}}}{\left(1+\mu\left(1-e^{-s}\right)\right)^N}$	$\mu_0+N\mu$	$N\mu(\mu+1)+(2\mu+1)\mu_0$
Continuous	$p_{\tilde{x}}(\tilde{x})$ ($\tilde{x}\ge 0$)	$M_{\tilde{x}}\left(e^{-s}\right)$	$\langle\tilde{x}\rangle$	$\sigma_{\tilde{x}}^2$
Constant	$\delta(\tilde{x}-\tilde{\mu}_0)$	$e^{-\tilde{\mu}_0 s}$	$\tilde{\mu}_0$	0
Central χ_{2N}^2, Gamma	$\frac{1}{\Gamma(N)}\tilde{\mu}^{-N}\tilde{x}^{N-1} \cdot e^{-\tilde{x}/\tilde{\mu}}$	$\frac{1}{(1+\tilde{\mu}s)^N}$	$N\tilde{\mu}$	$N\tilde{\mu}^2$
Non-central χ_{2N}^2	$\frac{\tilde{x}^{(N-1)/2}e^{-(\tilde{\mu}_0+\tilde{x})/\tilde{\mu}}}{\tilde{\mu}_0^{(N-1)/2}\tilde{\mu}} \cdot I_{N-1}\left(2\sqrt{\tilde{x}\tilde{\mu}_0}/\tilde{\mu}\right)$	$\frac{e^{\frac{-\tilde{\mu}_0 s}{1+\tilde{\mu}s}}}{(1+\tilde{\mu}s)^N}$	$\tilde{\mu}_0+N\tilde{\mu}$	$N\tilde{\mu}^2+2\tilde{\mu}\tilde{\mu}_0$

$$\begin{aligned} \frac{\partial M_n\left(e^{-s},t\right)}{\partial t} &= \sum_{n=-\infty}^{\infty} e^{-sn}\frac{dP(n,t)}{dt} = \sum_{n=0}^{\infty} -e^{-sn} \\ &\quad \times (n(a+b)+c)\,P(n,t) \\ &\quad + \sum_{n=1}^{\infty} e^{-sn}\,((n-1)a+c)\,P(n-1,t) \\ &\quad + \sum_{n=-1}^{\infty} e^{-sn}\,(n+1)\,bP(n+1,t) \quad . \end{aligned} \tag{57}$$

In the last but one sum n is replaced by $n+1$. In the last sum n is replaced by $n-1$. Using $\frac{\partial M_n(e^{-s},t)}{\partial s} = -\sum_{n=-\infty}^{\infty} ne^{-sn}P(n,t)$ we obtain

$$\begin{aligned} \frac{\partial M_n\left(e^{-s},t\right)}{\partial t} &= c\left(e^{-s}-1\right)M_n\left(e^{-s},t\right) \\ &\quad - \left(a-be^{s}\right)\left(e^{-s}-1\right)\frac{\partial M_n\left(e^{-s},t\right)}{\partial s} \end{aligned} \tag{58}$$

with the solution, at the amplifier output after group delay t,

$$\begin{aligned} M_n\left(e^{-s},t\right) &= \left(1+\mu\left(1-e^{-s}\right)\right)^{-N} \\ &\quad \times M_n\left(1-\frac{G\left(1-e^{-s}\right)}{1+\mu\left(1-e^{-s}\right)},0\right). \end{aligned} \tag{59}$$

Here $N=c/a$ is the number of noisy modes, $G=G(t)=e^{(a-b)t}$ is the accumulated power gain, $\mu=n_{sp}(G-1)$ is the expectation value of noise photons per mode and $n_{sp}=a/(a-b)$ is the spontaneous emission factor.

At the amplifier input we assume a Poisson distribution with mean $\tilde{\mu}_0=\mu_0/G$ and $M_n(e^{-s},0)=e^{-\tilde{\mu}_0(1-e^{-s})}$. Using (59) we get

$$M_n\left(e^{-s},t\right)=\left(1+\mu\left(1-e^{-s}\right)\right)^{-N}e^{\frac{-\mu_0\left(1-e^{-s}\right)}{1+\mu\left(1-e^{-s}\right)}} \tag{60}$$

and a noncentral negative binomial distribution (Table II) where

$$\begin{aligned} L_n^{\alpha}(x) &= \frac{1}{n!}e^x x^{-\alpha}\frac{d^n}{dx^n}\left(e^{-x}x^{n+\alpha}\right) \\ &= \sum_{m=0}^{n}(-1)^m\begin{pmatrix} n+\alpha \\ n-m\end{pmatrix}\frac{x^m}{m!} \end{aligned} \tag{61}$$

is a Laguerre polynomial. For $\tilde{\mu}_0=0$ (no input signal) one obtains a central negative binomial distribution at the output.

Let us cascade two optical amplifiers 1, 2. For a Poisson distribution with mean $\tilde{\mu}_0$ at its input the MGF $M_{n,1}(e^{-s},t_1)$ behind amplifier 1 is obtained as (60) with $\mu=\mu_1$, $G=G_1$ and $\mu_0=\mu_{0,1}=\tilde{\mu}_0 G_1$. Under the assumption that there is only one optical filter, behind all amplifiers and directly in front of the receiver, the mode number *N* is identical for both amplifiers. The MGF at the output of amplifier 2 is

$$\begin{aligned} M_{n,2}\left(e^{-s},t_2+t_1\right) &= \left(1+\mu_2\left(1-e^{-s}\right)\right)^{-N} \\ &\times M_{n,1}\left(\frac{1+(\mu_2-G_2)\left(1-e^{-s}\right)}{1+\mu_2\left(1-e^{-s}\right)},t_1\right) \end{aligned}$$

with $N=c_i/a_i$, $n_{sp,i}=a_i/(a_i-b_i)$, $G_i=e^{(a_i-b_i)t_i}$,

$$\mu_i=n_{sp,i}\left(G_i-1\right). \tag{62}$$

This can be rewritten as (60) with $t=t_2+t_1$, $\mu=\mu_2+\mu_1 G_2$, $\mu_0=\mu_{0,1}G_2=\tilde{\mu}_0 G_1 G_2=\tilde{\mu}_0 G$. The input-referred number of noise photons per mode is $\tilde{\mu}=\frac{\mu_1 G_2+\mu_2}{G_1 G_2}=\tilde{\mu}_1+\frac{\tilde{\mu}_2}{G_1}$. By induction we can extend this to a larger amplifier cascade, $\tilde{\mu}=\sum_{i=1}^{n}\frac{\tilde{\mu}_i}{\prod_{k=1}^{i-1}G_k}$, like (45).

Now we know photoelectron distributions for transmitted ones and zeros, can determine the optimum decision threshold where they intercept and can calculate BER. It depends on $\tilde{\mu}$, *G*, *N* and $\tilde{\mu}_0=2P\tau/(hf)$ where *P* is the mean optical power and τ is the observation or integration time, usually a bit shorter than the symbol duration. For BER $=10^{-9}$, $N=1$, $G\to\infty$, $\tilde{\mu}=1$ a mean of $\tilde{\mu}_0/2=38$ signal photons are needed at the amplifier input during τ. More in detail, let us consider polarization-independent optical amplifier, bandpass filter with rectangular impulse response envelope of duration $\tau_o=1/B_o$ and optical bandwidth B_o, photodetector and baseband filter consisting of *M* impulses spaced by τ_o each (Fig. 4). $\tau=M\tau_o$ is less or equal to the symbol duration. A real receiver will have lowpass characteristic of the photoreceiver and a rounded optical bandpass filter impulse response envelope. At the decision instant the sum of $N = pM$ independent negative binomial distributions is taken where $p = 2$ is the number of polarizations. Adding independent RVs means we can multiply MGFs. Multiplication of *N* negative binomial MGFs with equal $\tilde{\mu}$, *G* (and originally $N = 1$) yields (60). Given that there are two quadratures in each mode there are $2N=2pM$ degrees-of-freedom (DOF). Signal-spontaneous and spontaneous-spontaneous beat noises are automatically taken into account, exactly for our model receiver (Fig. 4). F_{pnf} is not directly needed.

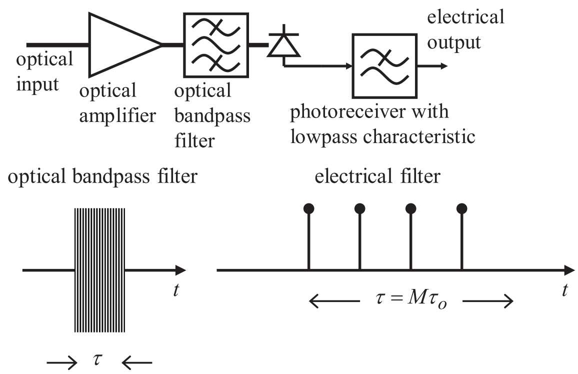

Fig. 4. Optical amplifier and receiver and impulse responses.

To understand optical amplifier noise better we consider $P(n)$ to be the result of the Poisson transform (25). $P(n)$ and $p_x(x)$ scale with *G*. No limit exists for MGF or distributions in the case $G\to\infty$. Yet *G* mainly scales the *x* range so that the PDF $p_{\tilde{x}}(\tilde{x})$ of a normalized input-referred intensity $\tilde{x}=x/G$ depends only weakly on *G*, and $\lim_{G\to\infty}p_{\tilde{x}}(\tilde{x})$ exists. Insertion of $x=G\tilde{x}$ into (25) results in a normalized Poisson transform

$$P(n)=\int_0^{\infty}p_{\tilde{x}}(\tilde{x})\,e^{-\tilde{x}G}\frac{(\tilde{x}G)^n}{n!}d\tilde{x} \tag{63}$$

with $p_{\tilde{x}}(\tilde{x})=Gp_x(G\tilde{x})$. After inserting (63) we obtain

$$\begin{aligned} \lim_{G\to\infty}M_n\left(e^{-s/G}\right) &= \lim_{G\to\infty}\sum_{n=0}^{\infty}P(n)\,e^{-(s/G)n} \\ &= M_{\tilde{x}}\left(e^{-s}\right). \end{aligned} \tag{64}$$

From that we can obtain $p_{\tilde{x}}(\tilde{x})$. In Table II, this converts a Poisson distribution into a Dirac distribution, i.e., constant intensity, and the negative binomial distributions become chi-squared distributions with 2*N* DOF. These are easier to evaluate. Among the variances of non-central RVs in Table II, the first summand corresponds to spontaneous-spontaneous beat noise, the second to signal-spontaneous beat noise.

Constant intensity yields a Poisson distribution due to shot noise. So, (64) eliminates shot noise by amplification and normalizes with respect to *G*. Inversely, (63) undoes the normalization and adds shot noise. $\tilde{x}$ and *x* contain optical amplifier noise (but no shot noise) and of course intensity modulation.

A chi-squared random variable is the sum of the squares of 2*N* independent Gaussian variables with equal variances. On the other hand, if we have a field (28) the (input-referred) intensity (30) is the sum of the squares of 2*N* independent Gaussian variables with equal variances $\tilde{\mu}/2$. In the absence of noise the expectation value is $\tilde{\mu}_0$, corresponding to the sum of squares of expectation values of fields. The only possible physical interpretation is: The optical amplifier adds Gaussian field noise in phase and in quadrature. The particle aspect is invoked upon detection: Photodetection adds shot noise.

For direct-detection DPSK and DQPSK receivers with interferometers, things are similar. At the interferometer outputs, differences of independent negative binomial (or chi-squared in the case $G \to \infty$) distributions with halved variances $\tilde{\mu}/4$ of Gaussian field noises are the decision variables. The MGF of such a difference $\tilde{x} = \tilde{x}_1 - \tilde{x}_2$ is $M_{\tilde{x}}(e^{-s}) = M_{\tilde{x}_1}(e^{-s}) M_{\tilde{x}_2}(e^{s})$. MGF, PDF and finally BER can be calculated. In some cases this even yields analytical BER results. For DPSK, $G \to \infty$, $N = 1$ a $BER = (1/2)e^{-\tilde{\mu}_0/\tilde{\mu}}$ is obtained.

For ASK the Gaussian approximation (taking means and variances from Table II and then calculating BER using the erfc function) is in reasonable agreement with the exact solution. This has contributed to the success of F_{pnf}. But for D(Q)PSK the Gaussian approximation yields BER considerably off the exact solution.

Coherent receivers with asynchronous detection for ASK, DPSK or DQPSK perform the same signal manipulations as their incoherent counterparts with 0, 1 or 2 interferometers. BER calculation is possible with the same mathematics and results. In coherent receivers there is usually only $p = 1$ active polarization.

Longer mathematical derivation, numerical results and diagrams are available for ASK in [7], [8] and for D(Q)PSK in [8].

Acknowledgment

The author would like to thank Peter Krummrich (TU Dortmund, Germany), for stimulating discussions.

References

[1] H. A. Haus, "IRE standards on methods of measuring noise in linear twoports, 1959," *Proc. IRE,* vol. 48, no. 1, pp. 60–68, Jan. 1960.
[2] H. T. Friis, "Noise figures of radio receivers," *Proc. IRE,* vol. 32, no. 7, pp. 419–422, Jul. 1944.
[3] E. Desurvire, *Erbium Doped Fiber Amplifiers: Principles and Applications*. Hoboken, NJ, USA: Wiley, 1994.
[4] R. Noé, *Optical Communication*. Paderborn, Germany: Paderborn Univ., 1995.
[5] H. A. Haus, "The noise figure of optical amplifiers," *IEEE Photon. Technol. Lett.*, vol. 10, no. 11, pp. 1602–1604, Nov. 1998, doi: 10.1109/68.726763.
[6] H. A. Haus, "Noise figure definition valid from RF to optical frequencies," *IEEE J. Sel. Topics Quantum Electron.*, vol. 6, no. 2, pp. 240–247, Mar./Apr. 2000,.
[7] R. Noé, "Optical amplifier performance in digital optical communication systems," *Elect. Eng.*, vol. 83, pp. 15–20, 2001.
[8] R. Noé, *Essentials of Modern Optical Fiber Communication*. Berlin, Germany: Springer, 2010.
[9] R. Noé, "kT/C noise: Covariance matrix calculation of instationary noise in time-varying systems," *AEÜ-Int. J. Electron. Commun.*, vol. 55, no. 2, pp. 135–139, 2001.
[10] K. Shimoda, H. Takahasi, and C. H. Townes, "Fluctuations in amplifications of quanta with application to maser amplifiers," *J. Phys. Soc. Japan.*, vol. 12, pp. 686–700, 1957.
[11] B. Saleh, *Photoelectron Statistics*. Berlin, Germany: Springer-Verlag, 1978.
[12] J. G. Proakis, *Digital Communications*, 2nd ed. New York, NY, USA: McGraw-Hill, 1989.
[13] W. Imajuku and A. Takada, "Noise figure of phase-sensitive parametric amplifier using a Mach-Zehnder interferometer with lossy Kerr media and noisy pump," *IEEE J. Quantum Electron.*, vol. 39, no. 6, pp. 799–812, Jun. 2003.
[14] M. Asobe, T. Umeki, and O. Tadanaga, "Phase sensitive amplification with noise figure below the 3dB quantum limit using CW pumped PPLN waveguide," *Opt. Exp.*, vol. 20, no. 12, pp. 13164–13172, 2012.
[15] P. A. Andrekson and M. Karlsson, "Fiber-based phase-sensitive optical amplifiers and their applications," *Adv. Opt. Photon.*, vol. 12, pp. 367–428, 2020.
[16] W. Lee and E. Afshari, "Low-noise parametric resonant amplifier," *IEEE Trans. Circuits Syst. I: Regular Papers*, vol. 58, no. 3, pp. 479–492, Mar. 2011, doi: 10.1109/TCSI.2010.2072370.

Reinhold Noe was born in Germany, in 1960. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Technical University of Munich, Munich, Germany, in 1984 and 1987, respectively.

From 1987 to 1988, he was with Bellcore (now Ericsson/Telcordia), Red Bank, NJ, USA. In 1987 and from 1988 to 1992 he was with Siemens in Munich, Germany. Since 1992, he has been a Professor of optical communication and high-frequency engineering with Paderborn University, Paderborn, Germany.

He has authored or coauthored more than 300 papers. In 2006, he and colleagues published the first coherent synchronous QPSK transmission with standard DFB lasers, in 2007 with polarization multiplex and electronic polarization control. In 2008, he was a co-recipient of, together with Ulrich Rückert, the Innovation Prize of the Land North-Rhine Westphalia, in the category innovation. In 2010, he founded, together with Benjamin Koch, Novoptel GmbH.