optica

Lasers with distributed loss have a sublinear output power characteristic

Tobias S. Mansuripur,[1] Guy-Mael de Naurois,[2] Alexey Belyanin,[3] and Federico Capasso[2,*]

[1]Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

[2]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

[3]Department of Physics and Astronomy, Texas A&M University, College Station, Texas 77843, USA

*Corresponding author: capasso@seas.harvard.edu

Received 19 September 2014; revised 26 November 2014; accepted 26 November 2014 (Doc. ID 223479); published 15 January 2015

It is a generally accepted fact of laser physics that in a homogeneously broadened gain medium, above threshold the output power of the laser grows linearly with the pump power. The derivation requires only a few simple lines in laser textbooks, and the linear growth is a direct result of the fact that above threshold, the intracavity optical intensity will increase to the point that the gain is saturated to the level of the net loss—so-called gain pinning or clamping. Such a derivation, however, assumes that the mirror loss is distributed (the approximation of uniform gain saturation) which is only a good assumption for cavities whose end mirrors have reflectivities close to one. Furthermore, in gain media with a distributed loss there is a maximum achievable intracavity intensity that in turn limits the output power. We show that the approximation of uniform gain saturation leads to output powers that violate this limit. More generally, for lasers with low mirror reflectivities that also have distributed loss, we prove that the output power grows sublinearly with the pump power close to threshold. Furthermore, after threshold the output grows linearly, but with a slope efficiency that can be substantially smaller than predicted by the uniform gain saturation theory, with the largest deviation occurring for traveling-wave lasers and asymmetric Fabry–Perot lasers. These results are particularly applicable to semiconductor lasers, and specific applications to quantum cascade lasers are discussed. © 2015 Optical Society of America

OCIS codes: (140.3430) Laser theory; (140.0140) Lasers and laser optics; (140.3560) Lasers, ring; (140.5960) Semiconductor lasers; (140.5965) Semiconductor lasers, quantum cascade.

http://dx.doi.org/10.1364/OPTICA.2.000048

1. INTRODUCTION

One of the signature characteristics of laser action in homogeneously broadened media is the linear growth of the output power with the pump power above the lasing threshold. In textbook derivations [1–7] this behavior is seen to be the direct result of gain clamping: above threshold, a steady state can only be reached when the intracavity optical intensity is such that the rate of stimulated emission reduces the population inversion to its threshold value, thereby always pinning the gain to the level of the net loss. Such a derivation, however, assumes that the gain saturates uniformly (in other words, that the intensity is constant throughout the gain medium) which is only a good assumption for lasers with high-reflectivity end mirrors. Surprisingly, we will show that this assumption also leads to the correct output power in the absence of distributed loss, no matter how small the reflectivities of the end mirrors. However, the uniform gain saturation approximation leads to unphysical predictions for the output power in cavities with large out-coupling that also contain a distributed intrinsic loss. The key point is that an unsaturable distributed loss imposes a maximum attainable intracavity intensity (which increases with the pump power) at which the number of photons generated by the saturated gain per unit length is equal to the number of photons discarded per unit length by the loss [8–10]. The uniform gain saturation treatment does not properly account for this simple fact, causing it to overestimate the output power of the laser. We will apply the equations first popularized by Rigrod [9,11–15] to account for the nonuniformity of the gain saturation and the lumped mirror losses, and demonstrate that, in the presence of distributed loss, the output power grows sublinearly with the pump power after threshold. Furthermore, past threshold the output grows linearly, but with a slope efficiency that can be significantly less than researchers would expect from the uniform gain saturation theory. The deviation of the output power from the prediction of the uniform gain saturation theory is largest for traveling-wave lasers (which display the largest variation in intracavity intensity) and also very significant for highly asymmetric Fabry–Perot (FP) cavities (i.e., one mirror with reflectivity much larger than the other). In symmetric FP cavities the deviation is the smallest, but can still be significant for particularly large values of distributed loss and out-coupling.

Much work has been done on diode lasers (the prototypical example of a laser with low facet reflectivities) to understand the nonuniformity of the gain and intracavity intensity. In this context the nonuniformity is often referred to as (long-range) spatial hole burning (SHB), but we prefer to reserve the term SHB to describe only the (short-range) nonuniformity due to standing-wave effects. An early theoretical work to understand the nonuniformity was given in [16], and experiments have confirmed the expected longitudinal gain distribution [17,18]. The fact that a nonuniform intensity distribution, in the absence of distributed loss, does not affect the power output was noted in [19], and we will generalize this result. The impact of distributed loss on the slope efficiency for a few specific cases was well-illustrated both computationally and experimentally in [20], but the understanding of the effect was limited. Our goal is to provide a physical understanding of the sublinear power output, which is a simple consequence of the intracavity intensity limit imposed by the distributed loss.

In what follows, we begin by explaining the geometry of the ring and FP lasers under consideration and introduce the Rigrod equation for homogeneously broadened gain media. A physical picture is presented to understand the uniform gain approximation in which the lumped transmission losses at the mirrors are replaced with a distributed out-coupling loss. With this picture in mind, it is simple to understand why the uniform gain approximation must fail for lasers with distributed intrinsic loss. Then, taking into account nonuniform gain saturation, we first analyze the traveling-wave laser, which yields an explicit formula for the output power. We also draw general conclusions about the behavior of both symmetric and asymmetric FP lasers, and compare them to the ring. We provide simple formulas that allow one to determine when the nonuniform gain saturation treatment should be used, and to estimate the magnitude of the discrepancy between the two theories. Finally, we point out how these new results are used to optimize laser design, using the quantum cascade laser (QCL) as a model system.

A FP laser cavity is shown in Fig. 1(a), which comprises a gain medium of length L placed between two lossless mirrors with reflectivities R_1 and R_2 (and transmissivities $1 - R_1$ and $1 - R_2$). The facets of the gain medium itself do not reflect, either through the use of Brewster-cut facets or anti-reflection coatings. The intensity envelopes of the right- and left-propagating waves are $I^+(z)$ and $I^-(z)$. Following the notation of [12], we will use the dimensionless ratios $\beta^+(z) = I^+(z)/I_{\mathrm{sat}}$ and $\beta^-(z) = I^-(z)/I_{\mathrm{sat}}$, where I_{sat} is the saturation intensity of the medium. The unsaturated gain, which is typically proportional to the pump power, is g_0 per unit length, and the distributed intrinsic loss is α_0 per unit length. We take g_0 to be independent of position, which assumes uniform temperature and, for injection lasers, uniform current density [16,17]. In Fig. 1(b), the same gain medium is employed in a traveling-wave configuration by placing it within a ring cavity comprising two perfect reflectors and one out-coupling mirror of reflectivity R. We will consider the unidirectional mode for which only β^+ is nonzero (achieved in practice with an optical isolator).

For a homogeneously broadened gain medium, the intensity envelopes obey the equations [1,12–15,21]

$$\frac{1}{\beta^+}\frac{d\beta^+}{dz} = -\frac{1}{\beta^-}\frac{d\beta^-}{dz} = \frac{g_0}{1+\beta^+ + \beta^-} - \alpha_0. \qquad \textbf{(1)}$$

We will consider only the homogeneously broadened case here; special cases of inhomogeneous broadening are discussed in [9,21,22]. By dealing only with the intensity envelope of the field, population grating effects resulting from SHB are not taken into account, but this treatment is nevertheless known to yield results in good agreement with experiment [1]. Equation (1) applies strictly to a single transverse and longitudinal mode; in a multimode laser the shape of the gain spectrum becomes important as does the mutual cross saturation of the various frequency components. While SHB is known to lead to multimode operation in homogeneously broadened lasers [21], SHB is not as significant in ring and asymmetric FP lasers, which display predominantly traveling-wave rather than standing-wave character. It is worth noting that our main

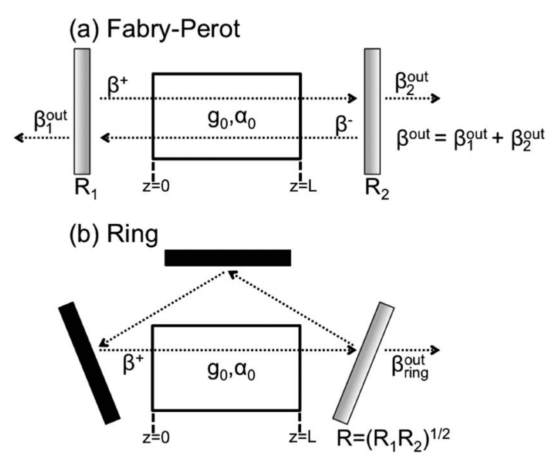

Fig. 1. Laser cavity geometry and relevant parameters: (a) FP laser and (b) unidirectional traveling-wave ring laser.

conclusion in this paper is most relevant to these two types of cavities where SHB is less important. Moreover, we expect our conclusions to still be valid in the multimode regime because the essential physics behind the results (namely, the intensity limit imposed by the distributed loss) is very general.

While Eq. (1), together with appropriate boundary conditions, is needed to solve for the laser output power above threshold, exactly at threshold the gain is completely unsaturated (i.e., intensity is zero) so the threshold condition is derived simply by demanding that the roundtrip unsaturated gain compensates both the intrinsic and out-coupling losses of the cavity. (Note that this derivation applies only when spontaneous emission is very weak, which is a good assumption for many kinds of lasers. A treatment that includes gain saturation at threshold due to spontaneous emission is given in [23,24]). To facilitate comparisons between the FP and ring lasers, we choose the reflectivity of the ring out-coupler to be $R = \sqrt{R_1 R_2}$ so that both cavities have the same threshold, given by

$$g_{0,\mathrm{th}} = \alpha_0 + \frac{1}{L} \ln\left(\frac{1}{R}\right). \tag{2}$$

2. DISTRIBUTED LOSS APPROXIMATION

The second term on the right-hand side of Eq. (2) is commonly called the mirror loss $\alpha_m = \ln(1/R)/L$, so that the total loss can be written as $\alpha_0 + \alpha_m$. We emphasize, however, that these two contributions to the loss behave very differently, α_0 being a distributed loss and α_m a lumped loss. A significant approximation made in textbook derivations of the laser output power [6,7], although not always explicitly stated, is to treat α_m as a distributed loss. In this way, all losses of the cavity become distributed, so we will refer to this as the distributed loss approximation (DLA). One way to think of the DLA is to imagine that the end mirrors of the cavity are replaced by perfect reflectors, and light escapes instead from any point within the gain medium with decay constant α_m. (Equivalently, a photon escape time τ_{esc} is often defined as $\tau_{\mathrm{esc}}^{-1} \equiv \alpha_m v_g$, where v_g is the group velocity of light in the cavity). Because the end mirrors have unity reflectivity, in the steady state the intracavity intensity does not vary with position; the saturated gain is exactly cancelled by the sum of the intrinsic loss and mirror loss at every position in the cavity. The DLA is therefore more commonly referred to as the uniform gain saturation approximation. Mathematically, Eq. (1) becomes

$$\frac{1}{\beta^+}\frac{d\beta^+}{dz} = \frac{g_0}{1+\beta} - (\alpha_0 + \alpha_m) = 0, \tag{3}$$

where $\beta \equiv \beta^+ + \beta^-$ is the total intensity. Equation (3) is easily solved for β and multiplied by the distributed out-coupling $\alpha_m L$ (as opposed to the lumped out-coupling $1 - R$, which is not the relevant out-coupling in the DLA) to get the total out-coupled intensity

$$\beta_{\mathrm{DLA}}^{\mathrm{out}} = \frac{\alpha_m L}{\alpha_0 + \alpha_m}(g_0 - g_{0,\mathrm{th}}). \tag{4}$$

Equation (4) is the well-known formula for the output power of a laser. In the DLA, both the FP and ring lasers emit the same intensity $\beta_{\mathrm{DLA}}^{\mathrm{out}}$. Note that the linearity of $\beta_{\mathrm{DLA}}^{\mathrm{out}}$ with the unsaturated gain g_0 (and therefore, in most cases, the pump power) follows from the DLA almost immediately.

There is a critical shortcoming of the DLA that is easy to understand. In a gain medium with intrinsic loss, whether an amplifier or a laser, there is a maximum attainable intracavity intensity $\beta^{\max}$ at which the number of photons generated by stimulated emission per unit length is exactly equal to the number of photons dissipated by the intrinsic loss per unit length [8–10]. By setting Eq. (1) to zero, we find that the total steady-state intracavity intensity, $\beta^+ + \beta^-$, cannot exceed

$$\beta^{\max} = \frac{1}{\alpha_0}(g_0 - \alpha_0), \tag{5}$$

at any position. This, in turn, places an upper limit on the out-coupled intensity. For simplicity, we consider the ring laser to explain the basic point: the output cannot exceed the mirror transmission times $\beta^{\max}$

$$\beta_{\mathrm{ring}}^{\mathrm{out,max}} = (1 - R)\beta^{\max} = \frac{1 - R}{\alpha_0}(g_0 - \alpha_0). \tag{6}$$

[Similar formulas apply to the FP laser in Eqs. (17) and (18).] It turns out that the DLA, by way of replacing the lumped mirror transmission $1 - R$ with the distributed out-coupling $\alpha_m L$, yields predictions for the output intensity that violate this maximum under certain conditions. Specifically, whenever the slope efficiency of $\beta_{\mathrm{DLA}}^{\mathrm{out}}$ is greater than the slope of $\beta_{\mathrm{ring}}^{\mathrm{out,max}}$,

$$\frac{\alpha_m L}{\alpha_0 + \alpha_m} > \frac{1 - R}{\alpha_0}, \tag{7}$$

the sketch in Fig. 2 illustrates that $\beta_{\mathrm{DLA}}^{\mathrm{out}}$ will exceed $\beta_{\mathrm{ring}}^{\mathrm{out,max}}$ above some value for g_0. In the next section we will solve explicitly for the output $\beta_{\mathrm{ring}}^{\mathrm{out}}$ taking into account nonuniform gain saturation, but we can already guess that the curve should resemble the dotted line in Fig. 2: it has the same threshold as the DLA predicts, but the output power must eventually be limited by $\beta_{\mathrm{ring}}^{\mathrm{out,max}}$. The parameter ranges of $\alpha_0 L$ and R that satisfy Inequality (7) are shown in light blue in the inset; in general, as $\alpha_0 L$ becomes larger and R smaller, the limiting influence of $\beta^{\max}$ becomes more relevant and nonuniform gain saturation must be accounted for. For example, for semiconductor lasers ($R \approx 0.3$), Fig. 2 indicates that the DLA is certainly problematic when $\alpha_0 L$ exceeds the value of 1.7, and we will show in subsequent sections that the discrepancy is measurable even for smaller values of $\alpha_0 L$.

3. NONUNIFORM GAIN SATURATION

We now address the behavior of the laser when the lumped nature of the mirror losses is properly accounted for, and

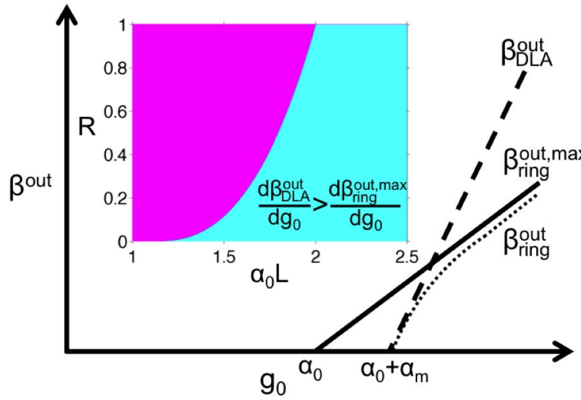

Fig. 2. Sketch illustrates that if the slope of $\beta^{\mathrm{out}}_{\mathrm{DLA}}$ (dashed) exceeds the slope of $\beta^{\mathrm{out,max}}_{\mathrm{ring}}$ (solid), the DLA prediction exceeds the allowable maximum output of the ring laser above a certain gain. Therefore, we can expect the actual output $\beta^{\mathrm{out}}_{\mathrm{ring}}$ to qualitatively follow the dotted curve. Inset, parameter ranges of $\alpha_0 L$ and R that result in such a violation, namely, those which satisfy Inequality (7), are shaded in light blue.

the gain saturates nonuniformly within the cavity. The general solution of Eq. (1) is given in [12], which we will restate here. It can be deduced from Eq. (1) that the product $\beta^+(z)\beta^-(z)$ is independent of z, so we designate the quantity $\beta_0^2 \equiv \beta^+(z)\beta^-(z)$. By making the substitution $\beta^-(z) = \beta_0^2/\beta^+(z)$, Eq. (1) is expressed entirely in terms of $\beta^+(z)$, which can then be integrated to yield

$$\alpha_0 L - \ln(\beta_1^+/\beta_2^+) = \frac{g_0 \ln[F(\beta_1^+)/F(\beta_2^+)]}{\sqrt{(g_0-\alpha_0)^2-(2\alpha_0\beta_0)^2}}, \tag{8}$$

where

$$F(\beta_i^+) \equiv \frac{\sqrt{(g_0-\alpha_0)^2-(2\alpha_0\beta_0)^2}+(g_0-\alpha_0-2\alpha_0\beta_i^+)}{\sqrt{(g_0-\alpha_0)^2-(2\alpha_0\beta_0)^2}-(g_0-\alpha_0-2\alpha_0\beta_i^+)}. \tag{9}$$

The subscript i on β^+ is either 1 or 2, and denotes evaluation of the intensity at the left facet ($z = 0$) or right facet ($z = L$), respectively. Note that Eqs. (8) and (9) apply to both the FP and ring lasers. For the FP, we use the boundary conditions at the mirrors

$$\beta_1^+ = R_1\beta_1^-, \qquad \beta_2^- = R_2\beta_2^+, \tag{10}$$

together with

$$\beta_0^2 = \beta_1^+\beta_1^- = \beta_2^+\beta_2^-, \tag{11}$$

to show that $\beta_1^+ = R\beta_2^+$ and $\beta_0 = \sqrt{R_2}\beta_2^+$, which allows us to express Eq. (8) entirely in terms of β_2^+. Still, in the general FP case Eq. (8) remains an implicit formula for β_2^+ which must be solved numerically. For the ring laser we substitute $\beta_1^+ = R\beta_2^+$ and $\beta_0 = 0$ into Eq. (8), which yields an analytic solution that will be discussed shortly.

In Supplement 1, we discuss two limiting cases for which the total output intensity as predicted by the nonuniform gain saturation theory in Eq. (8) reduces to that of the DLA: (1) small cavity out-coupling and arbitrary intrinsic loss α_0; and (2) arbitrary out-coupling and $\alpha_0 = 0$ (a surprising result, since the intensity variation in the cavity can be very large yet the DLA still predicts the correct output power [19]). In what follows, we will focus on the case of large out-coupling and large intrinsic loss, for which the predictions of the nonuniform gain saturation theory deviate significantly from those of the DLA.

A. Ring Laser

Unlike the FP, the equations for the unidirectional ring laser can be solved for an explicit expression of the output power. The existence of this solution (but not its expression) was stated in [12]. Starting from Eqs. (8) and (9) and using $\beta_0 = 0$ and $\beta_1^+ = R\beta_2^+$, we find

$$\beta^{\mathrm{out}}_{\mathrm{ring}} = (1-R)(g_0/\alpha_0 - 1) \times \frac{R\exp[(1-\alpha_0/g_0)(\alpha_0+\alpha_m)L]-1}{R\exp[(1-\alpha_0/g_0)(\alpha_0+\alpha_m)L]-R}. \tag{12}$$

We will explore various limits of Eq. (12). Close to threshold, the slope efficiency can be found by Taylor expanding the derivative of Eq. (12) with respect to g_0 to first-order in the parameter $(g_0 - g_{0,\mathrm{th}})/(\alpha_0+\alpha_m)$, which gives

$$\frac{d\beta^{\mathrm{out}}_{\mathrm{ring}}}{dg_0} = \frac{\alpha_m L}{\alpha_0+\alpha_m} - \frac{\alpha_0 L}{(\alpha_0+\alpha_m)^2}\left(\frac{1+R}{1-R}\ln(1/R)-2\right)(g_0-g_{0,\mathrm{th}}). \tag{13}$$

(We have also only retained terms to first-order in $\alpha_0 L$ for simplicity, although the following conclusions hold for large $\alpha_0 L$.) At threshold (i.e., $g_0 = g_{0,\mathrm{th}}$), the slope efficiency is given by the first term of Eq. (13), which is the same as the prediction of the DLA in Eq. (4). However, while the DLA predicts a constant slope efficiency above threshold, in fact the slope efficiency decreases as g_0 increases past $g_{0,\mathrm{th}}$ as given by the second term of Eq. (13). Note that the term in large parentheses is always positive for $0 < R < 1$.

As g_0 continues to increase, the slope efficiency asymptotically approaches a constant value, which we can see by Taylor expanding Eq. (12) in the high-gain limit $g_0 \gg g_{0,\mathrm{th}}$ (without making any assumption about $\alpha_0 L$), which gives

$$\beta^{\mathrm{out}}_{\mathrm{ring}}\Big|_{g\gg g_{0,\mathrm{th}}} = \frac{(1-R)(1-e^{-\alpha_0 L})}{\alpha_0(1-Re^{-\alpha_0 L})}(g_0 - g^{\mathrm{eff}}_{0,\mathrm{th}}), \tag{14}$$

where

$$g^{\mathrm{eff}}_{0,\mathrm{th}} \equiv (\alpha_0+\alpha_m)(\alpha_0 L)\left(\frac{1}{1-e^{-\alpha_0 L}} - \frac{1}{1-Re^{-\alpha_0 L}}\right) + \alpha_0, \tag{15}$$

is an effective threshold that would be found by linearly extrapolating the power output at high gain back to the g_0 axis. We will refer to the derivative of Eq. (14) with respect to g_0 as the asymptotic slope efficiency, which is clearly independent of

g_0. In the limit of moderate to large intrinsic loss, $\exp(\alpha_0 L) \gg 1$, and arbitrary R, Eq. (14) reduces to

$$\beta_{\text{ring}}^{\text{out}}\big|_{g \gg g_{0,\text{th}}\ \&\ \exp(\alpha_0 L) \gg 1} = \frac{1-R}{\alpha_0}(g_0 - \alpha_0), \tag{16}$$

which is recognizable as $\beta_{\text{ring}}^{\text{out,max}}$ of Eq. (6). Thus, we see that for moderate to large $\alpha_0 L$, at high gain the intracavity intensity reaches its limiting value $\beta^{\max}$, so the output intensity follows $\beta_{\text{ring}}^{\text{out,max}}$, as illustrated by the dotted line in Fig. 2. For smaller $\alpha_0 L$ that does not satisfy $\exp(\alpha_0 L) \gg 1$, the output intensity requires the full Eq. (14); in this case the intracavity intensity never grows sufficiently to approach $\beta^{\max}$, but the asymptotic slope efficiency is nevertheless smaller than the DLA slope efficiency.

To understand quantitatively how much the asymptotic slope efficiency differs from the near-threshold slope efficiency, let us look at the ratio of the slope efficiency in the high-gain limit [slope of Eq. (14)] to the slope efficiency at threshold [the first term of Eq. (13), which is the same as the slope efficiency predicted by the DLA]. This ratio is plotted in Fig. 3 as a function of R for five different values of the product $\alpha_0 L$. As expected, the discrepancy between the asymptotic and threshold slope efficiency is greatest for small R and large $\alpha_0 L$. The effect of $\alpha_0 L$ can be understood with recourse to Fig. 2: as α_0 increases the slope of $\beta_{\text{ring}}^{\text{out,max}}$ decreases, and as L increases the slope of $\beta_{\text{DLA}}^{\text{out}}$ increases. In either case, an increase in the product $\alpha_0 L$ widens the disparity between the asymptotic and threshold slope efficiency.

Finally, while the asymptotic output intensity was derived in the limit $g_0 \gg g_{0,\text{th}}$, it turns out that this is not a very stringent limit. For a full quantitative analysis, see Supplement 1. Here we emphasize the conclusion of that analysis: the full solution in Eq. (12) closely approaches the asymptotic solution in Eq. (14) when g_0 is only a few multiples of $g_{0,\text{th}}$, and often when g_0 is even smaller. This is significant because in some lasers, such as QCLs, g_0 can never exceed more than a few multiples of $g_{0,\text{th}}$, and so this asymptotic regime is still experimentally accessible. Second, as $\alpha_0 L$ is increased at fixed R, the asymptotic regime is reached at ever smaller values of g_0. This trend can also be understood easily with recourse to Fig. 2: note that the gain-intercept of $\beta_{\text{ring}}^{\text{out,max}}$ is α_0 while the laser threshold is $\alpha_0 + \alpha_m$. As $\alpha_0 L$ increases at fixed R, the intercept α_0 approaches $\alpha_0 + \alpha_m$, and so the laser must enter the asymptotic regime sooner after threshold. Let us consider one numerical example: for $\alpha_0 L = 10$ and $R = 0.25$, the asymptotic slope efficiency is 62% of the threshold slope efficiency (see Fig. 3), and the asymptotic regime is reached once $g_0 \geq 1.4 g_{0,\text{th}}$ (Supplement 1). This example will be extended to the FP laser in the next section.

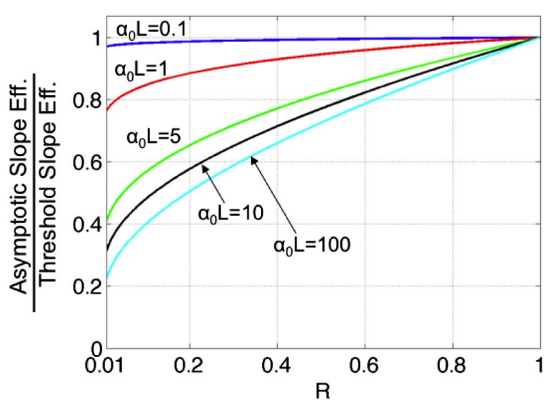

Fig. 3. Ring laser: ratio of the asymptotic to the threshold slope efficiency as a function of R for various values of $\alpha_0 L$.

B. Fabry–Perot Laser

The maximum output intensity of the ring laser $\beta_{\text{ring}}^{\text{out,max}}$ in Eq. (6) was derived by assuming that the intracavity intensity reaches $\beta^{\max}$ at the out-coupling mirror. Similarly, the maximum output intensities of the symmetric and maximally asymmetric FP lasers are

$$\beta_{\text{sFP}}^{\text{out,max}} = 2\left(\frac{1-R}{1+R}\right)\beta^{\max}, \tag{17}$$

$$\beta_{\text{aFP}}^{\text{out,max}} = \left(\frac{1-R^2}{1+R^2}\right)\beta^{\max}. \tag{18}$$

For the maximally asymmetric FP, the reflectivity of the left mirror is unity and the other is $R_2 = R^2$. The denominators $1 + R$ and $1 + R^2$ in Eqs. (17) and (18), respectively, arise from the need to transform the total intensity $\beta^{\max}$ into the right-traveling component β_2^+, which is then multiplied by the transmissivity $1 - R$ (and a factor of 2 for the two mirrors) or $1 - R^2$, respectively. Inequalities similar to Inequality (7) can easily be written down using Eqs. (17) and (18), which then serve as useful indicators for when to be wary of the DLA solution for FP lasers.

Since we cannot analytically examine the output power of the FP laser as we did for the ring laser, we consider a concrete example for which $\alpha_0 = 20\ \text{cm}^{-1}$, $L = 0.5$ cm, and $R = 0.25$, and solve numerically for the output power. (These are representative values for a QCL with emission wavelength between 8 and 12 μm [25]). In Fig. 4(a), the total output intensity β^{out} as a function of g_0 is shown for three lasers: (1) symmetric FP with $R_1 = R_2 = 0.25$, (2) maximally asymmetric FP with $R_1 = 1$ and $R_2 = 0.25^2 = 0.0625$, and (3) ring laser with $R = 0.25$. According to the DLA, all three lasers should have the same output intensity, and this curve is plotted as well. The maximum output intensity for each laser, given by Eqs. (6), (17), and (18), is plotted in dashed lines. In the inset of Fig. 4(a), the slope efficiency of each laser is shown alongside the derivative of $\beta^{\text{out,max}}$; for every laser, the slope of $\beta^{\text{out,max}}$ is less than the slope efficiency of the DLA, which means the curve $\beta^{\text{out,max}}$ will fall beneath $\beta_{\text{DLA}}^{\text{out}}$ above some value of g_0. This is clearly seen in Fig. 4(a) for the ring and asymmetric FP, and occurs outside the plotted domain (at $g_0 = 212\ \text{cm}^{-1}$) for the symmetric FP. We see that all three lasers have the same slope efficiency at threshold, and the value is correctly predicted by the DLA. (For the ring laser we already proved this fact in the previous section). As the gain increases, the slope efficiencies of all three lasers decrease, as they must because once the asymptotic regime is reached the slope

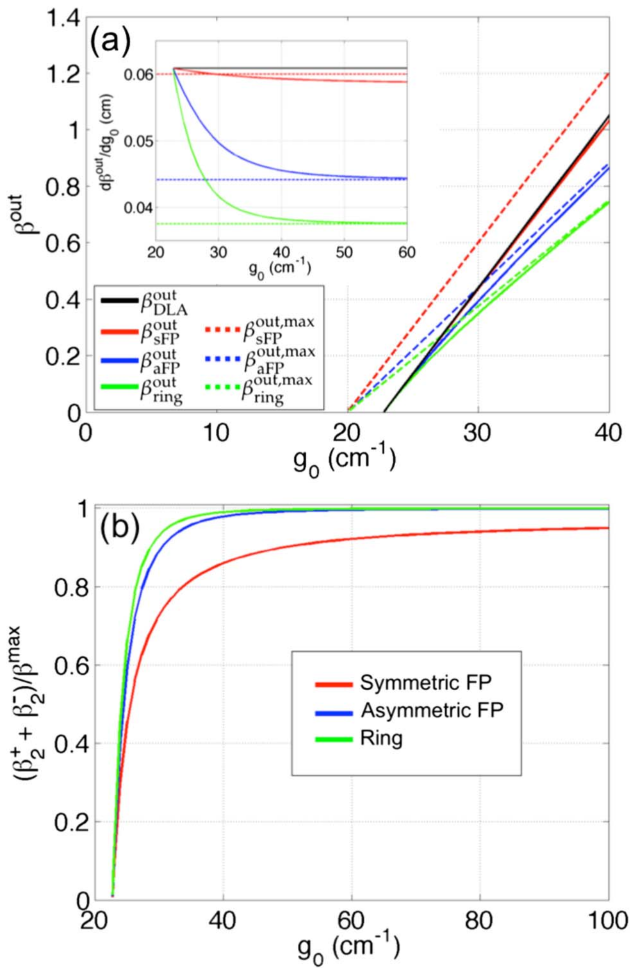

Fig. 4. (a) β^{out} versus g_0 for $\alpha_0 = 20\ \text{cm}^{-1}$, $L = 5$ mm, and $R = 0.25$, shown for the DLA (black), symmetric FP with $R_1 = R_2 = 0.25$ (red), maximally asymmetric FP with $R_1 = 1$ and $R_2 = 0.25^2 = 0.0625$ (blue), and ring laser with $R = 0.25$ (green). The maximum output intensity is shown in correspondingly colored dashed lines. Inset, slope efficiencies of the curves (i.e., first derivatives), together with the corresponding slopes of $\beta^{\text{out,max}}$ for each laser; (b) intracavity intensity at the right facet, $\beta_2^+ + \beta_2^-$, is normalized to $\beta^{\max}$ and plotted against g_0 over a larger range than in (a).

efficiency must be less than the slope of $\beta^{\text{out,max}}$. In the asymptotic regime, the symmetric FP has the largest slope efficiency (96% of threshold value), followed by the asymmetric FP (72% of threshold value) and then the ring (62% of threshold value). In this example, the DLA is not a bad approximation for the symmetric FP because $\beta_{\text{sFP}}^{\text{out,max}}$ does not violate $\beta_{\text{DLA}}^{\text{out}}$ over a large range of gain. However, the DLA does not provide accurate output intensities for the asymmetric FP and ring, for which the limiting effects of $\beta_{\text{aFP}}^{\text{out,max}}$ and $\beta_{\text{ring}}^{\text{out,max}}$ are stronger and occur soon after threshold.

Another way to understand the output power of the lasers in Fig. 4(a) is shown in Fig. 4(b): here, the intracavity intensity at the right mirror (where it is a maximum), $\beta_2^+ + \beta_2^-$, is normalized to $\beta^{\max}$ and plotted as a function of g_0. Note that the normalization factor $\beta^{\max}$ is itself an increasing function of g_0. In the ring laser and asymmetric FP, shortly after threshold the total intensity at the right mirror quickly approaches $\beta^{\max}$. To be exact, in the asymptotic regime the intracavity intensities approach $0.99997\beta^{\max}$ and $0.99987\beta^{\max}$ in the ring and asymmetric FP, respectively. In the symmetric FP, the total intensity at the right mirror asymptotically approaches the slightly smaller value of $0.974\beta^{\max}$, although for larger values of $\alpha_0 L$ it can be shown to also approach $\beta^{\max}$. Thus, for sufficiently large $\alpha_0 L$, Eqs. (6), (17), and (18) do not provide merely an upper limit, but rather a good approximation of the output intensity in the asymptotic regime. What do we mean by sufficiently large $\alpha_0 L$? For the ring laser we previously showed the necessary condition to be $\exp(\alpha_0 L) \gg 1$, but for the FP this is not easy to answer without numerical calculation. As a general trend, however, the ring laser always has the largest intracavity intensity at the out-coupling mirror, followed by the asymmetric FP and then the symmetric FP. Thus, the symmetric FP requires the largest amount of loss $\alpha_0 L$ in order for the intensity to reach $\beta^{\max}$, followed by the asymmetric FP and then the ring. To explain this, it helps to plot the intracavity intensity along the entire length of the cavity, as shown in Fig. 5 for the same laser parameters as used in Fig. 4. In the symmetric FP, the left- and right-propagating waves have the most similar intensities at each position; the cross saturation of the two waves hinders the growth of each and prevents the total intracavity intensity from reaching $\beta^{\max}$. In the asymmetric FP, β^+ (which travels from the HR mirror to the lower-reflectivity mirror) is much more intense than β^-. Therefore, β^- does not cross-saturate β^+ significantly, which allows β^+ to grow quickly. As a result, we see that the total intensity closely approaches $\beta^{\max}$. In the ring laser, the absence of β^- allows β^+ to grow even more quickly at small z and approach the limit $\beta^{\max}$ well before it reaches the right mirror (see Supplement 1 for a mathematical discussion of the intensity variation in the various cavities). The key point is that cross-saturation in the symmetric FP reduces the total intracavity intensity. However, for large enough $\alpha_0 L$, even this mechanism does not prevent the intensity from reaching $\beta^{\max}$.

4. OTHER APPLICATIONS

The DLA overestimates the increase in output power that can be achieved by increasing the length of the laser. As a practical example, we consider an asymmetric FP QCL with $\alpha_0 = 15\ \text{cm}^{-1}$, high-reflectivity (HR) coating on one facet $R_1 = 1$, and low-reflectivity (LR) coating on the other $R_2 = 0.01$ (an achievable value with current mid-infrared coating technology) to increase the single-ended output power. In practice every laser has a maximum achievable gain g_0, so suppose our goal is to maximize the output power at $g_0 = 40\ \text{cm}^{-1}$. In Fig. 6, the output power β^{out} vs. g_0 is plotted for two such lasers, one with $L = 4$ mm and one with $L = 5$ mm, for both the DLA and nonuniform gain saturation theory. In the DLA, the longer laser outputs 12% more power at $g_0 = 40\ \text{cm}^{-1}$ ($\beta^{\text{out}} = 2.40$ instead of 2.13). In reality, however, the longer laser can be expected to output only 3% more power ($\beta^{\text{out}} = 1.59$ instead of 1.54). The reason for this small increase is that the 4 mm laser is already emitting close to the maximum possible intensity $\beta_{\text{asym}}^{\text{out,max}}$ of Eq. (18), also plotted in Fig. 6, so increasing the length of the laser will

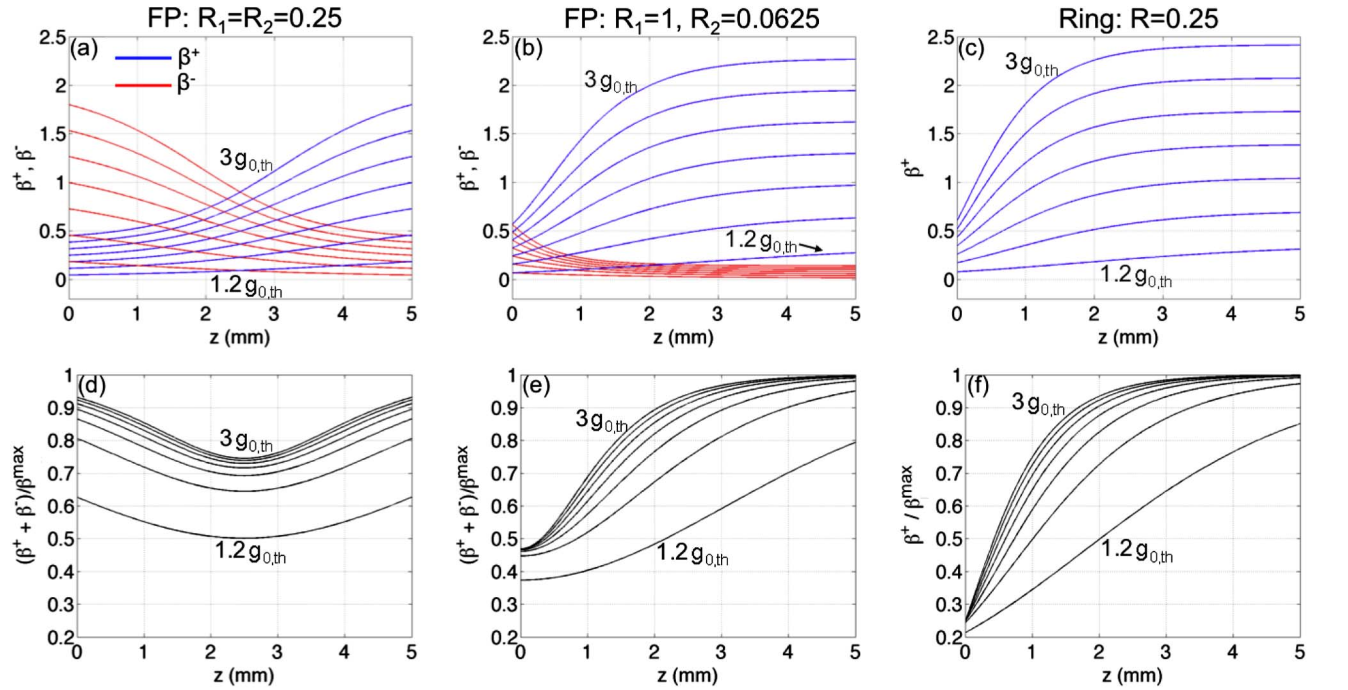

Fig. 5. For the same parameters as those used in Fig. 4 ($\alpha_0 = 20\ \mathrm{cm}^{-1}$, $L = 5$ mm, and $R = 0.25$), the intensity envelopes β^+ and β^- of the right- and left-propagating waves are plotted for $1.2g_{0,\mathrm{th}}$ to $3g_{0,\mathrm{th}}$ (in steps of $0.3g_{0,\mathrm{th}}$); (a) for a symmetric FP with $R_1 = R_2 = 0.25$; (b) for a maximally asymmetric FP with $R_1 = 1$ and $R_2 = 0.25^2 = 0.0625$; (c) for a ring laser with $R = 0.25$. Total intracavity intensity, $\beta^+ + \beta^-$, is normalized to $\beta^{\max}$ and plotted in (d), (e), and (f) for the same three lasers, respectively. See Supplement 1 for a mathematical discussion of the intensity variation.

not increase the output significantly. In this example, a 25% increase in the length of the cavity results in only a 3% increase in power output, which clearly leads to a significant decrease in the wall-plug efficiency. More generally, for lasers with large intrinsic loss there is an optimum length beyond which the wall-plug efficiency will decrease.

The output intensity of a laser is maximized for a particular end mirror reflectivity, and this optimal reflectivity will be significantly miscalculated when using the DLA. Note that Rigrod provides analytic formulas that account for nonuniform gain saturation to approximate the optimal reflectivity [12], but the formulas become less accurate for lasers with large intracavity intensity variation. This is the case for the asymmetric FP and ring, so here we will calculate the optimum numerically. As before, let us consider $\alpha_0 = 15\ \mathrm{cm}^{-1}$, $R_1 = 1$, and $L = 4$ mm, though now our goal is to find the optimum reflectivity R_2 that maximizes the power output at $g_0 = 40\ \mathrm{cm}^{-1}$. The nonuniform gain saturation theory predicts an optimum reflectivity $R_2 = 5.4 \times 10^{-3}$ leading to $\beta^{\mathrm{out}} = 1.54$. The DLA predicts an optimum reflectivity $R_2 = 5.0 \times 10^{-4}$, nearly 1 order of magnitude smaller. Using a coating with the DLA optimum reflectivity would result in $\beta^{\mathrm{out}} = 1.50$, which is only a 2.6% reduction in output power relative to the true optimum. This is a small difference, but the importance of this example is that the optimal coating reflectivity is 1 order of magnitude larger than predicted by the DLA. This is important for active research in LR mid-infrared coatings: for the purpose of maximizing laser output power, we do not need to aim for reflectivities as low as we had thought.

One final application concerns laser characterization. The ratio of the experimentally measured slope efficiencies of two lasers, identical except for their facet coatings, whose reflectivities are known, can be used to determine the value of α_0 by comparison with the DLA slope efficiency formula [26]. In such cases, the result can be skewed if one of the lasers is highly asymmetric, and a preferable method to determine the distributed loss would be to vary the cavity length L while maintaining the cavity symmetry $R_1 = R_2$ (or use the threshold values

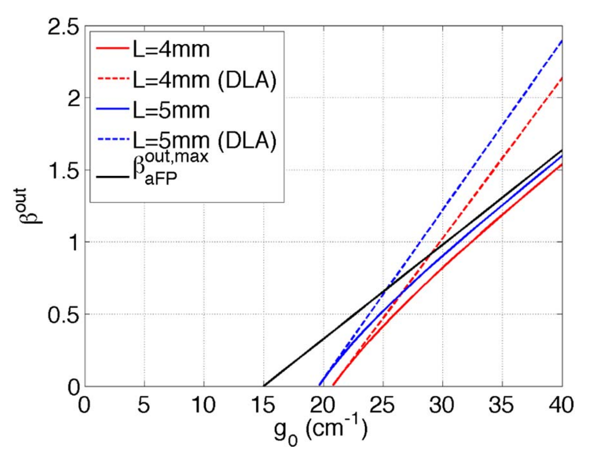

Fig. 6. Output power $\beta^{\mathrm{out}}_{\mathrm{aFP}}$ of the maximally asymmetric FP laser with $\alpha_0 = 15\ \mathrm{cm}^{-1}$, $R_1 = 1$, and $R_2 = 0.01$, for cavities of length 4 mm (red) and 5 mm (blue), as computed numerically accounting for nonuniform gain saturation (solid lines) and as predicted by the DLA (dashed lines). Maximum output power $\beta^{\mathrm{out,max}}_{\mathrm{aFP}}$ is also plotted (black).

rather than the slope efficiencies, which are unaffected by the degree of symmetry of the cavity). Note that the results of [26] are not affected by this phenomenon because in their case one of the lasers had a HR facet and a bare facet, as opposed to a LR facet. Based on numerical calculations we have done, this would not have been enough asymmetry to affect their results by an experimentally significant margin.

5. CONCLUSION

The uniform gain saturation treatment of a laser effectively replaces the transmission of light at the end mirrors with an out-coupling mechanism distributed throughout the cavity. We referred to this as the DLA, and showed that it is a good approximation to the nonuniform gain saturation treatment when the out-coupling is small or when the intrinsic loss of the gain medium is zero. When the intrinsic loss is not zero, there is a maximum attainable intracavity intensity $\beta^{\max}$ at which the number of photons generated by stimulated emission per unit length is equal to the number of photons absorbed or scattered by the intrinsic loss per unit length. When the intensity in the laser nears this maximum, the DLA overestimates the power that can be coupled out of the cavity. The result is that the output power grows sublinearly with the pump power, in stark contrast to the famous linear relationship derived in the uniform gain saturation approximation. We provided simple formulas in terms of $\alpha_0 L$ and R (applicable to all ring and FP lasers) to help determine the magnitude of the output power discrepancy between the DLA and the nonuniform treatment. The deviation is greatest for traveling-wave lasers and highly asymmetric FP lasers.

While our derivation explicitly treats only single-mode lasers and neglects SHB, in fact we expect the sublinear power output to be a quite general effect. To be completely rigorous, the standing-wave effects of SHB should be accounted for, and in a multimode laser the gain cross-saturation of each frequency on the others complicates the mathematics significantly. Therefore, the value of $\beta^{\max}$ for each mode will be different and difficult to calculate. Nevertheless, the limiting effect of $\beta^{\max}$ will be small near threshold and substantial far above threshold, as we have shown. This is the essential ingredient for a sublinear power output.

Our assumption of uniform small-signal gain g_0 along the length of the laser is a common one. In electrically injected lasers, this assumption relies on a uniform current injection. This is not necessarily the case in highly efficient lasers, if the electrical resistance of the laser decreases significantly with an increasing rate of stimulated emission. In such a laser, the current density would concentrate in the regions where the gain is most highly saturated. This effectively delivers electrons to where they are needed most, and would oppose the limiting influence of $\beta^{\max}$. One could also intentionally deliver more current to the region of highest intensity using multiple electrical contacts [20], or taper the waveguide width to dilute the optical intensity and prevent it from nearing $\beta^{\max}$ [20]. Both strategies would mitigate the sublinear power output.

FUNDING INFORMATION

NSF (ECCS-1230477).

See Supplement 1 for supporting content.

REFERENCES

1. A. Siegman, *Lasers* (University Science Books, 1986).
2. O. Svelto, *Principles of Lasers*, 5th Ed. (Springer, 2010).
3. W. T. Silfvast, *Laser Fundamentals*, 2nd Ed. (Cambridge University, 2004).
4. A. Yariv and P. Yeh, *Photonics: Optical Electronics in Modern Communications*, 6th Ed. (Oxford University, 2007).
5. B. E. Saleh and M. C. Teich, *Fundamentals of Photonics*, 2nd Ed. (Wiley, 2007).
6. S. L. Chuang, *Physics of Photonic Devices*, 2nd Ed. (Wiley, 2009).
7. L. A. Coldren and S. W. Corzine, *Diode Lasers and Photonic Integrated Circuits* (Wiley, 1995).
8. E. Schulz-DuBois, "Pulse sharpening and gain saturation in traveling-wave masers," Bell Syst. Tech. J. **43**, 625–658 (1964).
9. W. Rigrod, "Gain saturation and output power of optical masers," J. Appl. Phys. **34**, 2602–2609 (1963).
10. N. Karlov, *Lectures on Quantum Electronics*, 2nd Ed. (CRC, 1993).
11. W. Rigrod, "Saturation effects in high-gain lasers," J. Appl. Phys. **36**, 2487–2490 (1965).
12. W. Rigrod, "Homogeneously broadened CW lasers with uniform distributed loss," IEEE J. Quantum Electron. **14**, 377–381 (1978).
13. G. M. Schindler, "Optimum output efficiency of homogeneously broadened lasers with constant loss," IEEE J. Quantum Electron. **16**, 546–549 (1980).
14. T. R. Ferguson, "Lasers with saturable gain and distributed loss," Appl. Opt. **26**, 2522–2527 (1987).
15. T. R. Ferguson and W. P. Latham, "Efficiency and equivalence of homogeneously broadened lossy lasers," Appl. Opt. **31**, 4113–4121 (1992).
16. S. Hasuo and T. Ohmi, "Spatial distribution of the light intensity in the injection lasers," Jpn. J. Appl. Phys. **13**, 1429–1434 (1974).
17. W. W. Fang, C. Bethea, Y. Chen, and S. L. Chuang, "Longitudinal spatial inhomogeneities in high-power semiconductor lasers," IEEE J. Sel. Top. Quantum Electron. **1**, 117–128 (1995).
18. F. Rinner, J. Rogg, P. Friedmann, M. Mikulla, G. Weimann, and R. Poprawe, "Longitudinal carrier density measurement of high power broad area laser diodes," Appl. Phys. Lett. **80**, 19–21 (2002).
19. B. Ryvkin and E. Avrutin, "Spatial hole burning in high-power edge-emitting lasers: A simple analytical model and the effect on laser performance," J. Appl. Phys. **109**, 043101 (2011).
20. Z. Chen, L. Bao, J. Bai, M. Grimshaw, R. Martinsen, M. DeVito, J. Haden, and P. Leisher, "Performance limitation and mitigation of longitudinal spatial hole burning in high-power diode lasers," Proc. SPIE **8277**, 82771J (2012).
21. L. W. Casperson, "Laser power calculations: sources of error," Appl. Opt. **19**, 422–434 (1980).
22. A. White, E. Gordon, and J. Rigden, "Output power of the 6328-Å gas maser," Appl. Phys. Lett. **2**, 91–93 (1963).
23. L. W. Casperson, "Threshold characteristics of multimode laser oscillators," J. Appl. Phys. **46**, 5194–5201 (1975).
24. L. W. Casperson and M. Khoshnevissan, "Threshold characteristics of multimode semiconductor lasers," J. Appl. Phys. **75**, 737–747 (1994).
25. C. Gmachl, F. Capasso, D. L. Sivco, and A. Y. Cho, "Recent progress in quantum cascade lasers and applications," Rep. Prog. Phys. **64**, 1533–1601 (2001).
26. A. Wittmann, T. Gresch, E. Gini, L. Hvozdara, N. Hoyler, M. Giovannini, and J. Faist, "High-performance bound-to-continuum quantum-cascade lasers for broad-gain applications," IEEE J. Quantum Electron. **44**, 36–40 (2008).