Quantum Cascade Lasers – 25 Years After the First Demonstration

C. Kumar N. Patel, Mariano Troccoli, and Rodolfo Barron-Jimenez
Pranalytica, Inc., 1101 Colorado Avenue, Santa Monica, CA 90401

1. INTRODUCTION

This year marks the 25th anniversary of the first demonstration of quantum cascade lasers by Faist, et al [1]. QCLs have proven to be unique sources of mid wave infrared (MWIR) and long wave infrared (LWIR) radiation, capable of operating in continuous wave mode at room temperature. QCLs are the only solid state sources for generating laser radiation at these wavelengths, converting electrical power directly into optical energy. The only other sources that meet the criterion of direct conversion of electrical power into optical power are the molecular gas lasers, such as carbon dioxide [2, 3] and carbon monoxide [4], which were first demonstrated fifty five years ago. Both of these systems have been commercially developed to produce high powers necessary for many industrial, defense and medical applications. However, both the CO_2 and the CO laser operate on multiple discrete infrared transitions of the respective molecules and thus are only discretely tunable over limited wavelength ranges, from ~9.0 μm to ~11.5 μm for the CO_2 laser (using normal and isotopic CO_2) and from ~5.0 μm to ~7.0 μm for the CO laser. On the other hand, QCLs are able to cover a spectral range from ~3.5 μm to terahertz, *albeit,* with much lower power. In spite of the lower powers available from QCLs, there are a very broad range of applications that have made QCLs attractive for continued research and development.

QCLs are unipolar semiconductor lasers that operate on distinct energy levels created within the conduction band of the semiconductor material and are unlike the traditional semiconductor lasers that operate on the gain spectrum created by band-to-band recombination radiation in a junction between n and p type materials. Efficient operation of QCLs requires efficient injection of electrons into the upper laser level (see Figure 1). These electrons will fall to the lower laser level providing gain at the frequency corresponding to the energy separation between the upper and lower laser levels. For continued operation of the laser, the lower laser level needs to be depopulation efficiently. Both of these processes are reminiscent of gas lasers. However, unlike the gas lasers the lasing level lifetimes of much shorter, typically of the order of picoseconds, leading to substantial broadening of the emission line.

The discrete energy levels arise from confinement of conduction band electrons in potential wells created by the growth of ultrathin (~nm) layers of two semiconducting materials with different bandgaps. At present only the system comprising of alternate layers of $In_xAl_{1-x}As$ and $In_yGa_{1-y}As$ has been shown to be usable for QCLs, where $In_xAl_{1-x}As$ has a bandgap larger than that of $In_yGa_{1-y}As$. This leads to the generation of quantized electron levels within the conduction band offset of the two materials, which are utilized for gain generation as described above. The layers' thickness and the conduction band offset between the two materials determine the band structure of the

discrete quantized levels (subbands), thus providing a way to determine the energy separation between the upper and lower laser levels, and as a consequence the emission wavelength of the device, by design.

Each well/barrier active multilayer region is interleaved with injection/relaxation regions that have the twofold purpose of enhancing the upper laser level lifetime by Bragg confinement and reducing the lower laser level lifetime by resonant optical phonon scattering. This way, the designed energy band structure can support the buildup of population inversion between the upper and lower energy levels of the laser transition during current injection. Alternatively repeating the active and injection multilayers creates a sort of cascade effect for the electrons that are extracted from one stage and injected into the next. This structure can be repeated several tens of times until enough gain volume is provided to sustain laser action.

a) Single LO transition   b) Multi-phonon relaxation   c) Non-resonant extraction

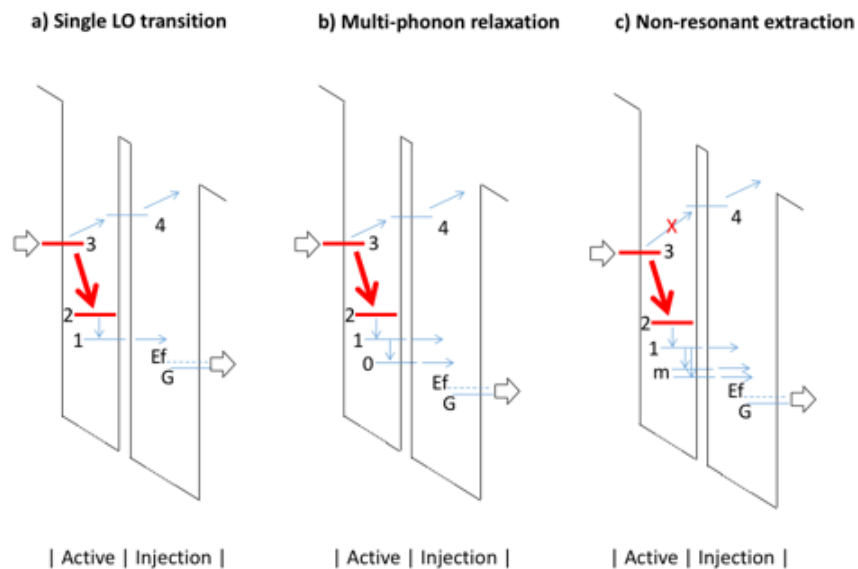| Active | Injection | | Active | Injection | | Active | Injection |

Figure 1: schematic examples of a single active / injection period for different QCL band structure designs. The red lines and arrow indicate the laser levels and transition. (a) traditional single-optical phonon relaxation "2-1", (b) double optical phonon relaxation "2-1-0" (c) non-resonant extraction to multiple states "2-1-m" with optimized current injection by reduced leakage "3-4".

Efficient current injection into the upper energy state of the laser transition is one of the key qualities of high efficiency design. Under the appropriate voltage bias, electrons populate the lowest energy states of the injection/relaxation regions with a Fermi distribution that to a first approximation we can assume to be thermalized at the lattice temperature. The highest momentum electrons from the reservoir so created will undergo resonant tunneling through the AlInAs injection barrier into the upper laser level of the next stage and thus contribute to the buildup of population inversion. Nevertheless, there are a number of other energy states in the active material and there is a non-zero total probability of electrons scattering to states other than the upper laser

level. An optimized injection design will minimize this probability and therefore have minimal current "leakage" to non-lasing states of the active wells. As a consequence, a highly efficient injection design will also be less sensitive to temperature variations and therefore have a higher T_0 and T_1, leading to higher power output at room temperature.

Besides efficient current injection, the second element leading to population inversion and laser action is the fast depletion of the lower laser level. This can be achieved by designing the active region in a 3- or 4-level scheme where electrons in the lower laser level have at least one further energy step that is resonant with the energy of a longitudinal-optical (LO) phonon. The strong coupling of the electron-LO phonon scattering at resonance will significantly reduce the lifetime of the lower laser level compared to the upper laser level and lead to population inversion and lasing. This concept is illustrated in Figure 1a.

In single-phonon relaxation designs, at room temperature the thermal population of optical phonons will induce a non-negligible probability for re-absorption of the phonons by the low-energy electrons and therefore produce a back-filling effect of the lower laser level with consequent reduction if not altogether extinction of the population inversion. Therefore, for high temperature operation the need for multiple phonon relaxation steps and a deeper built-in potential has been key to high performance, room temperature operation of QCLs. This concept is illustrated in Fig 1b. [5]

Lastly, the most recent evolution of the relaxation scheme is the so-called "non-resonant extraction", where the electron energy relaxation via single optical phonon is replaced by multiple transition channels created in the injection/relaxation region via a miniband of states that have an overall comparable scattering rate to the LO-phonon resonant scattering. This way, the designer has an added flexibility in the choice of active region layers without them being "locked" to the LO phonon energy for the transitions to the lower energy states. This concept is illustrated in Fig 1c. [6]

2. CURRENT STATUS AND PERFORMANCE OF QCLs

QCLs are intrinsically high power devices, due to the high field / high current densities that are typical of their operation characteristics. Because of the intrinsically short intersubband transition lifetimes, typical operating currents densities will necessarily be in the range of 1 kA/cm^2 as illustrated by a simple estimate:

$$J = q\ N_D\ /\ \tau \sim 10^3\ A/cm^2 \qquad (1)$$

where q is the electron charge, $N_D \sim 10^{10}$ cm^{-2} is the typical QCL doping sheet density and $\tau \sim$ 1ps is a typical intersubband transition lifetime. This is one of the main reasons why thermal management at room temperature has been one of the main drivers to improve QCL performances. On the other hand, the high current densities can be exploited to generate high power beams if losses and current leakage can be minimized

by design as discussed in the previous section. From the very first demonstration of CW room temperature operation in 2001, where only about 17mW were observed, [5] today we have demonstrated multi-Watt emission from QCLs at room temperature. [6]

Highest power output from the QCLs are obtained when the laser is operated in a Fabry-Perot configuration, with rear facel high-reflection coated and the from facet coated for controlled antireflection coated for maximum power extraction. In this configuration, the laser output covers a broad spectrum, reflecting the broad gain distribution arising from the unintentional heterogeneous broadening from the center wavelengths of each of the 40-50 gain stages not being identical. Typical spectral width of the output spectrum is about 100 cm^{-1}.

One of the key elements determining the device emission characteristics is the gain bandwidth. Typical QCL devices are based on the repetition of identical active regions that have the same laser transition by design. The intrinsic linewidth broadening of such transition depends on a number of factors, including carrier-carrier scattering, interface scattering, material quality, among others. Typical broadening of the laser linewidth is in the range of ~14meV for high quality material, and can get worse (up to ~30meV) for lower quality material. A narrow linewidth generally corresponds to better laser performance as fewer of the photons generated by the laser transition are "wasted" in spontaneous emission and the lasing threshold can be reached at lower currents.

State-of-the-art performances for high power QCLs (FP configuration) are presently in the range of several Watts per single emitter across the whole "sweet spot" for QCL operation, i.e. the mid-IR region from λ=4.0μm up to λ=9μm. Beyond this range, some fundamental limitations of the InP-based material start to significantly affect laser efficiency and high power performance.
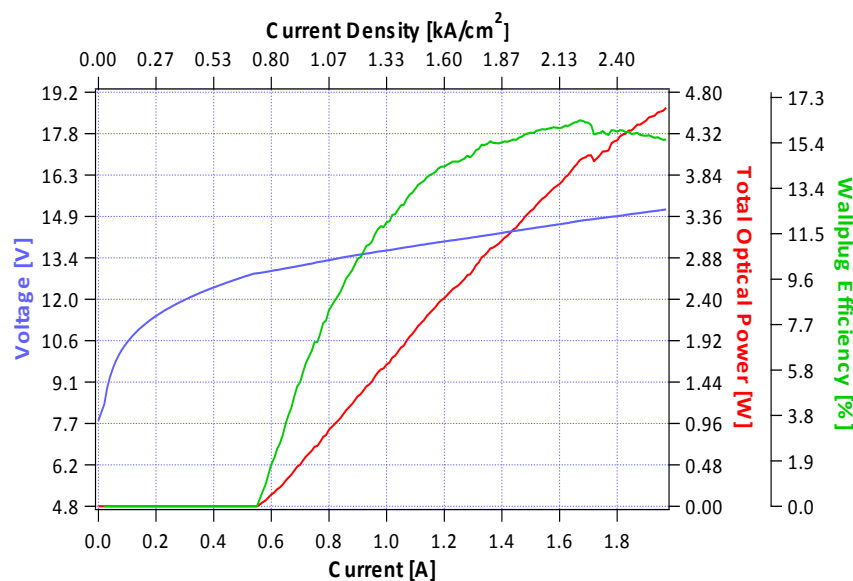

Figure 2: output characteristics of a 4.6um QCL with 10mm cavity length and 7.5um ridge width. A total of 4.5W of power was recorded corresponding to more than 16% WPE (Ref. 5).

Pranalytica has shown some of the best performances across the spectrum, both for pulsed and CW / QCW bias conditions. QCW average output powers of up to P=4W and conversion efficiencies of WPE=16% have been obtained in the λ=4.6μm region, while up to P=2W and WPE=10% have been recorded in the λ=9μm region. In the strategically important λ=4.0 μm region, the best performances have been shown with a power output of P=2W and a WPE exceeding 5%. The record performances are illustrated in Figures 2, 3, and 4.
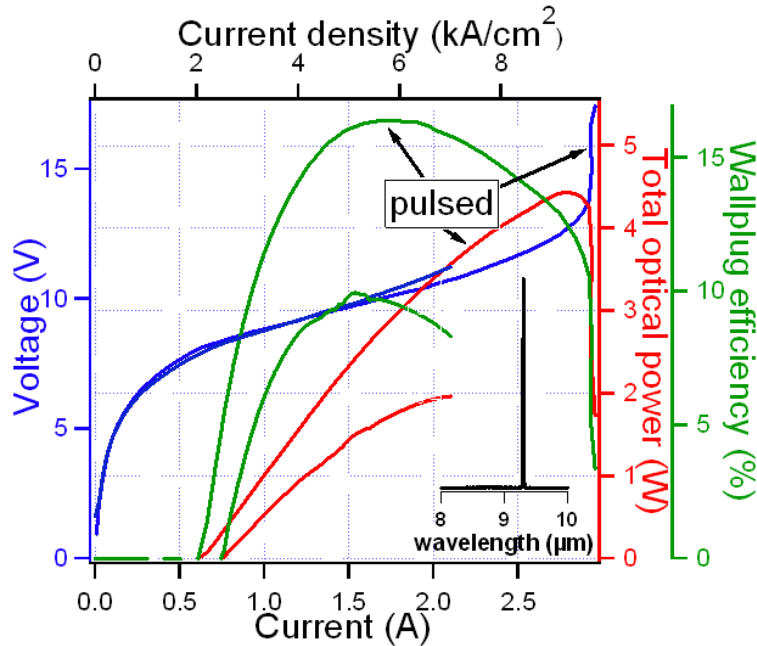

Figure 3: Comparison between pulsed and CW total optical power vs current and voltage vs current characteristics measured at 293K for an uncoated 3mm by 10µm laser mounted epi-down on AlN/SiC composite submount. Inset shows pulsed laser spectrum taken at maximum current.
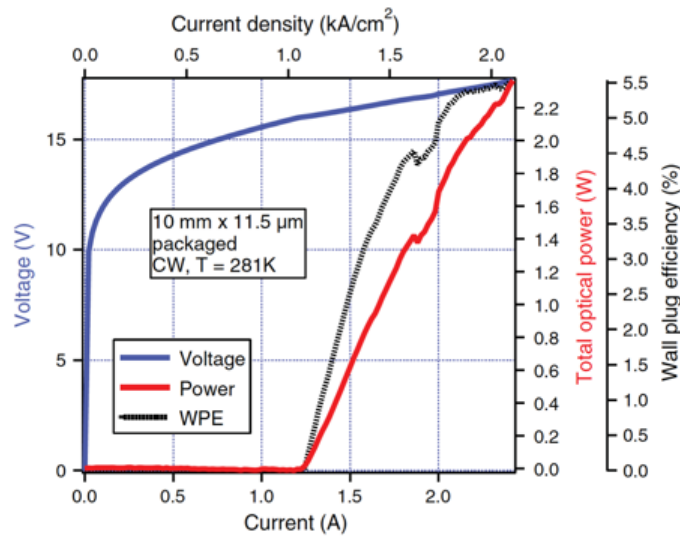

Figure 4: Output characteristics of a λ=4.0μm QCL with 10mm cavity length and 11.5um ridge width at room temperature. A maximum emission power of 2W and a WPE of 5% are recorded.

Tables 1 and 2, below summarize maximum continuous wave and average power output (both at room temperature) obtained from FP quantum cascade lasers in CW and QCW (high duty cycle pulsed operation) modes respectively. In CW mode, the QCLs require thermoelectric cooler for thermal management. In QCW mode, the QCLs can be operated without TEC because when operating in this mode, typically with 500 ns pulses at a pulse repetition rate of ~1 MHz, there is enough time available between consecutive pulses to remove the heat from the QCL active region with QCL epi-side down mounted on the submount. The data presented in Tables 1 and 2 are for QCLs that are in production at present (not "hero" experimental results). For Table 2, the system efficiency represents the conversion of DC electrical power to the pulse driver electronics to optical power.

Table 1. Typical best continuous wave power output at four sample wavelengths achieved to date from FP configuration single emitter quantum cascade lasers operating in CW mode

Highest Performance CW/RT QCLs in Production (w/TEC)		
Wavelength	**CW Power Output (W)**	**WPE**
~4.6 µm	> 3.5 W	> 16%
~4.0 µm	> 2.0 W	~ 10%
~7.1 µm	> 1.4 W	~ 10 %
~9.3 µm	> 2.0 W	~ 10 %

Table 2. Typical best average power output at four sample wavelengths achieved to date from FP configuration single emitter quantum cascade lasers operating in QCW mode

Highest Performance QCW/RT QCLs in Production (No TEC)		
Wavelength	**Average Power Output (W)**	**System Efficiency**
~4.6 µm	> 3.0 W	> 10%
~4.0 µm	> 2.0 W	~ 7 %
~7.1 µm	> 1.5 W	~ 12 %
~9.3 µm	> 2.0 W	~ 8 %

Thermal management in high power quantum cascade lasers

As can be seen from the two tables above, the electrical power to optical power conversion efficiency is between ~10% to 16%. Thus, for example for a 3 W average power output from a QCL at 4.6 µm (in QCW mode), the laser generates ~27 W of heat in the active region, which needs to be removed. Thus, the submount must have high thermal conductivity. Moreover, production version QCLs are mounted epi-side down on the submounts with hard solder (not In solder) for long term reliability, which

necessitates a good thermal expansion match between the QCL material and the submount. Typically, AlN is used as a submount material, which has been extensively used in InP based semiconductor lasers. AlN has good thermal conductivity (typically ~100 W/m-K) and a good coefficient of thermal expansion (CTE) match with the QCL material. However, we have developed a new submount material, which is a hybrid of AlN and SiC. Silicon carbide has a thermal conductivity of ~500 W/m-K, but its CTE match with QCL material is poor. Using the AlN/SiC composite material, we take advantage of the high thermal conductivity of SiC and the good CTE match of AlN [6,7].

Optimization of QCL performance by reducing the number of gain stages: Typical quantum cascade laser design involves a structure that has 40-50 gain stages, which has the advantage that it provides enough gain to overcome the cavity losses and a gain region thickness (~2 μm), which is proportional to number of stages, sufficient to assure a large mode overlap factor, Γ, with the active region. However, the thermal management requires removal of heat from the active layers even when the QCL is mounted epi-side down on the submount material. The primary path for heat removal is across the epi layers, a direction in which the thermal conductivity of the material is poor. Thus, having 40-50 gain stages provides sufficient gain and volume for mode containment, but could result higher active region temperature. Recently, we have explored the possibility of using fewer than 40-50 gain stages. Simulations showed that the reduction in gain (and therefore the maximum power output) can be compensated by increasing injected power density. Because of the thickness of the epi layer stack is proportional to the number of gain stages, the increased heat generation in the active region does not lead to an increase in the active region temperature because of the more efficient thermal transport through a shorter distance. In fact the active region temperature, even with substantially increased level of injected power, should be lower than that with 40 gain stages because improved thermal transport. The table below shows the results of simulation for 40, 30 and 20 gain stages.

Table 3. Simulated performance of QCLs with reduced number of gain stages

Number of Stages	Active Region Thickness (μm)	Electrical Power Density (10^{14} W/m^3)	T_{max} (K)	T_{av} (K)
40	1.85	2.6	425	374
30	1.35	3.56	403	364
20	0.9	5.34	377	352

The simulation data show that in spite of increasing the electrical power density to compensate for the gain and power reduction, active region temperature actually is lower for fewer number of gain stages.

Experimentally we have verified the performance for a 20 gain stage QCL [8] as seen in Figure 5. Lyakh and his colleagues have explored this concept further by showing that use of smaller number of gain stages and a broader mesa for the QCL still results in a high beam quality output with higher power output [9].
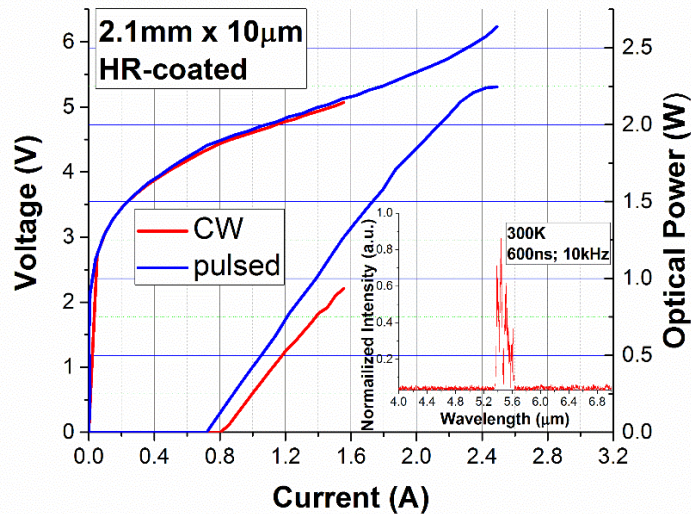
Figure 5. Pulsed (600 ns; 10 kHz) and CW optical power vs. current and voltage vs current characteristics measured for a HR-coated 2 mm °— 10 μm QCL at 293 K. The inset shows that the laser spectrum is centered at 5.55 μm [7].

Narrow gain spectrum versus broad gain spectrum

While in general, it is desirable to have the output spectrum of a FP configuration QCL be as narrow as possible, i.e., to have the gain center for each of the 40-50 gain stages to be identical and provide no heterogeneous gain broadening in addition to the homogeneous broadening arising from the natural lifetimes of the lasing levels, there are some cases where emission over a broad spectral range is desirable. In particular, tunable lasers, utilizing external optical cavity that includes a spectral selection element such as a mechanically rotated diffraction grating or an electronically tuned acousto-optic grating for obtaining wide tuning, can benefit from a broad gain medium.

The strength of the QCL structure is that it can be designed and tailored to the specific application needs. In the case of broad gain, a "heterogeneously" broadened active gain material can be obtained by design via the careful assembly of multiple slightly different active regions in the same device core. While a typical QCL with a single active will be able to tune only ~100 cm^{-1}, the heterogeneous broadening approach can lead to several 100's of wavenumbers of spectral coverage across the mid-IR with a single emitter in an appropriately designed external cavity.
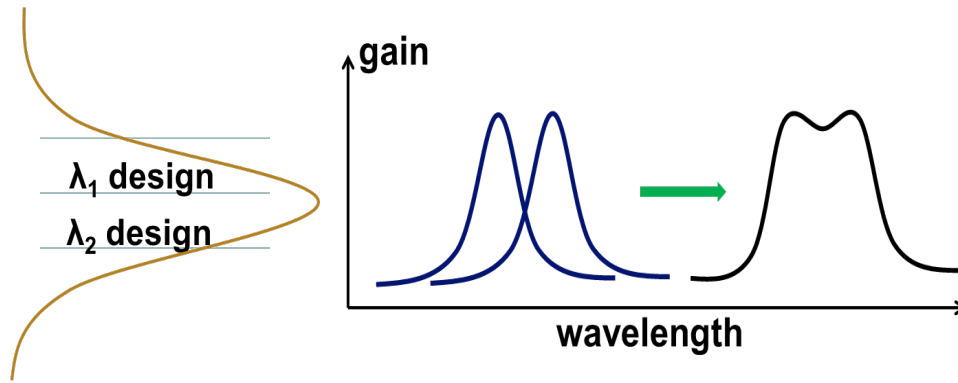
Figure 6. Schematics of heterogeneous design approach (left) and resultant gain spectral profile (right)

Frequency combs in QCLs

A second mode of operation where broad spectral coverage is desirable is frequency combs (FC). QCL combs do not require optical feedback with sensitive alignment characteristic such as, for example, external cavity lasers. As a consequence, they can be ruggedized for employment in the most demanding applications.

In QCL frequency combs, material gain spectral width by use of heterogeneous QCLs is on the order of several hundred wavenumbers. As a consequence, the active medium can support up to one thousand longitudinal optical modes separated by approximately $0.3\text{-}0.5\,cm^{-1}$. The best design techniques developed for high power/high efficiency broadly tunable QCLs can be extended to the QCL comb structure design, provided that the overall spectral gain/loss profile is structured to result into a wide zero-dispersion spectral region, a condition conducive to mode locking via four-wave mixing (FWM) in the active region.

By carefully choosing a number of active region stages for each design in the heterogeneous stack and their central emission wavelengths, an almost arbitrary spectral gain profile can be engineered. Since gain and refractive index are related through the Kramers-Kronig relations, the control of gain profile allows for the engineering of refractive index dispersion. In particular, it is possible to design a heterogeneous QCL structure with a nearly zero dispersion in a wide spectral region. This method was used for the original demonstration of frequency combs [10] and it was subsequently used to further improve device performance [11].

The intermode spacing in a QCL FC is defined by device cavity length and its typical value is approximately $0.5\ cm^{-1}$. The spacing may be too large for some spectroscopic applications when the FC modes do not coincide with targeted absorption lines. In this case, the spectral position of FC lines can be tuned by changing laser temperature: Similar to DFB QCLs, mode position can be changed by approximately $1\ cm^{-1}$ when device temperature changes on the order of 10K.

A powerful spectroscopic technique is based on the use of two combs from identical lasers slightly detuned so that the spacing between the modes of the two combs increases gradually across the spectrum. By observing the beating signal between the modes the spectral signatures can be converted into an RF spectrum where each frequency is directly correlated to one position in the optical spectrum. A simultaneous measurement of the whole spectral range can thus be achieved in an extremely short time without the need for a wavelength sweeping mechanism.

3. TUNABLE QUANTUM CASCADE LASERS

In the above discussions of performance of QCLs, we have focused on FP configuration of the optical cavity, which provides the highest amount of optical power output, *albeit* at a fixed QCL wavelength determined primarily by the QCL structure design. As with all lasers that have broad gain profile, there is additional value of the broad gain medium, which is capability of generating tunable radiation of substantially narrower bandwidth than that is obtained from FP configuration QCLs. Such output is of substantial value in spectroscopy and spectroscopy based applications in the MWIR and the LWIR spectral regions, since these regions cover the strong fundamental optical absorptions of very large number of gases, liquids and solids. As alluded to earlier in this paper, tunable output QCL systems are configured using external cavities (rather than the optical cavity formed by the end facets of the QCLs), which include appropriate spectral selection element. The simplest of this is the one in which one of the mirrors of the external cavity configuration is replaced by a diffraction grating that can be mechanically rotated to provide a wavelength selective resonant optical cavity (Figure 7).
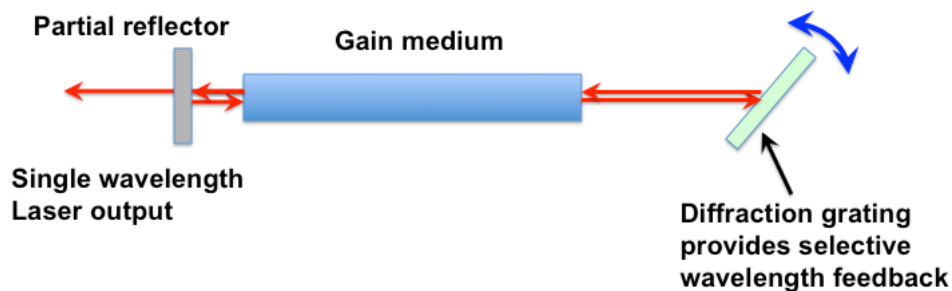

Figure 7. Sketch of an external grating cavity QCL configuration (Littrow geometry)

In general, the configuration shown in Figure 7, permits full exploitation of the spectral gain features of the QCL gain medium (as long as the gain exceed the cavity losses). The speed with which the wavelengths can be changed is limited by the mechanical mass of the diffraction grating. In this configuration, it is possible to speed up the tuning rate by keeping the diffraction grating in a fixed angular location and including an additional reflecting element such as a micro-electro-mechanical system (MEMS), which can provide rapid change in the resonant wavelength of the optical cavity.

Recently, we have introduced yet another way of obtaining tunable radiation from QCL gain element [12]. Instead of using the diffraction grating for providing spectral selection, we include an acousto-optical element that can provide a phase grating in responses to an applied RF (Figure 8). As can be seen, there are no movable parts and all the optical elements are permanently fixed in their locations. Light travelling through the acousto-optic modulator (AOM) is refracted by the phase grating set up within the AOM by the Bragg angles that depends only on the materials properties of the AOM (refractive index and the acoustic velocity), the RF drive frequency and the laser frequency. Keeping the high reflectivity mirror fixed in space as well as in angle, a change in the RF drive frequency changes the spatial wavelength of the phase grating in the AOM and thus a different laser wavelength is now resonant as defined by Equation 2 below.

$$\sin\theta_B = \frac{m\,n\,\lambda_o}{\Lambda} \tag{2}$$

where m is the order of refraction, n is the refractive index of the AOM material, λ_o is the optical wavelength and Λ is the acoustic wavelength in the AOM. For an optical wavelength, λ_0 = 10 μm, refractive index of ~4.0 for the germanium AOM crystal, and an acoustic wavelength of ~135 μm (assume an acoustic driving frequency of 35 MHz), we get a diffraction angle of θ_B = ~7.5°.
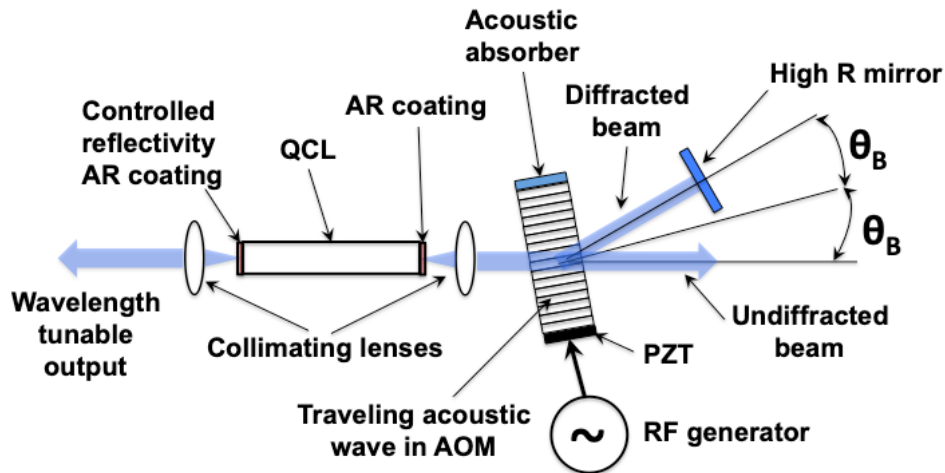

Figure 8. Acousto-optic modulator tuned quantum cascade laser system

In a demonstration experiment, where the QCL chip had a gain profile that covered a spectral region from ~8.2 μm to ~9.8 μm, application of RF from 43.5 MHz to 46.5 MHz laser tuning covered a spectral range from 1020 cm^{-1} (9.8 μm) to 1170 cm^{-1} (8.5 μm). Since the acoustic wavelength in the AOM can be changed very rapidly by changing the RF drive frequency, laser wavelength tuning can be equally rapid. The laser wavelength switching speed is independent of the size of the wavelength change and depends only the acoustic wave transit time across the laser beam diameter within the

AOM material. Measured wavelength switching speed of ~500 ns is in agreement with the theoretical expectations. For the example cited above, the measured switching speed corresponds to a scan of ~150 cm^{-1} in 0.5 μs, yielding scan speed of ~3 x 10^8 cm^{-1}/s!

As a test of the scanning speed, we measured the infrared absorption of R134A CFC (Figure 9), which shows the expected features of the spectrum, obtained in a single "shot" of ~17 μs, the measurement speed being limited by the response time of the optical detector.

Finally, AOM tuned QCL system permits simultaneous operation on two independently controlled wavelengths by an application of two different RF frequencies. Each wavelength can be tuned and switched independently by controlling the respective RF [8]. This capability makes it possible to carry out pump-probe studies of materials in the MWIR/LWIR region and also provides a tool for estimating intrinsic QCL lifetimes [13].

Unlike the traditional grating tuned QCL systems (described earlier), the AOM tuned QCL systems do not require a different diffraction grating for use in a different spectral scan region. For example going from 10 μm to 5 μm, will require only a different QCL chip and a change in the angle θ_B at which the totally reflecting mirror is located and then permanently fixed.
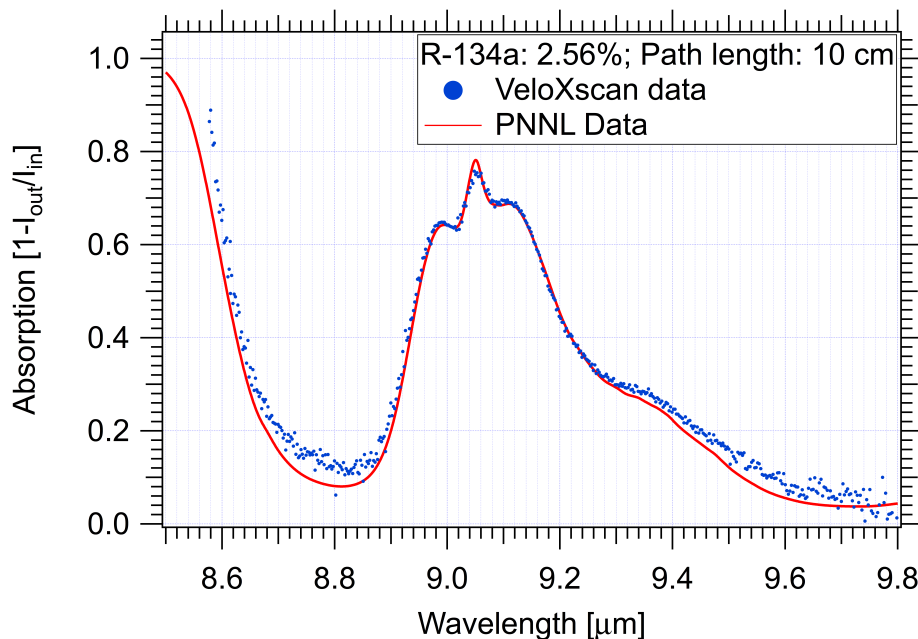

Figure 9. Absorption spectrum of R134a CFC measured in a single shot scan of 17 μs, using the tunable radiation from the AOM tuned QCL system,

The AOM tuned QCL system has been named VeloXscan and can be seen in Figure 10.

Figure 10. AOM tuned QCL system (VeloXscan) laser head

4. QCL PACKAGING INNOVATIONS

High power FP configuration QCLs can be packaged very compactly, as seen from one of the configurations that weighs <50 g and has a volume of ~2 in^3 (Figure 11).
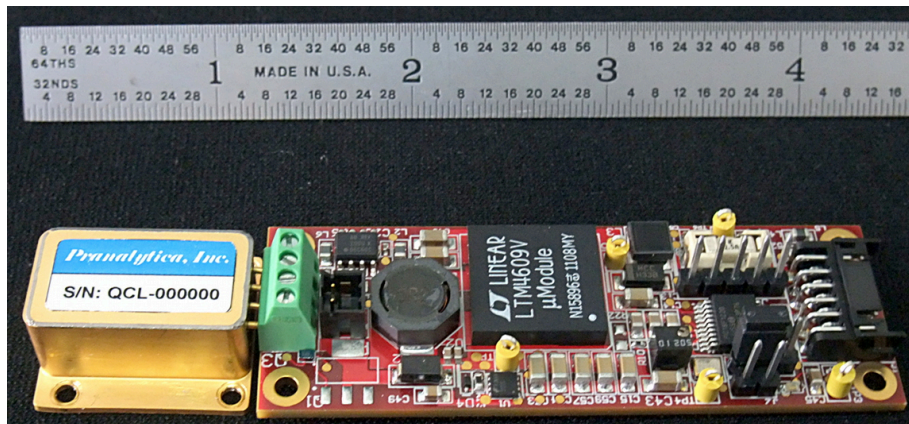

Figure 11. Compact package for high average power output QCW operation quantum cascade lasers (for performance see Table 2)

The butterfly package (left in Figure 10) houses the FP QCL chip and beam collimating optics, together with a thermistor to monitor the QCL operating temperature. The butterfly package is hermetically sealed for long term reliability. The electronics driver board is powered by 12-24VDC and provides the high current drive at 500 ns pulse width with a pulse repetition rate of ~1 MHz to deliver high average power output at the sample wavelengths shown in Table 2. This laser configuration has been tested to meet MIL-STD 518G vibration, mechanical shock and thermal shock standards. Because of the very small weight and size (as well as the high power and high electrical

to optical power conversion efficiency, Table 2), the package sown in Figure 11 is ideal for hand held, battery operated applications, such as target pointers and IFF beacons.

Often, there is need for two laser wavelengths from a single system that produces high power at the two wavelengths, which can be independently controlled. Traditional way to accomplish this is to use two separate QCL packages of the type shown in Figure 10 and dichroically combine the radiation at the two wavelengths using mirrors and other optical elements mounted on adjustable mirror mounts. While this arrangement works well, it is suited only for laboratory environment where one does not encounter vibration and mechanical as wells as thermal shocks. Such an arrangement is hardly suitable for portable or field deployable application, where size and weight as well as ruggedness are critical.

To overcome the above limitations, we have evolved the single QCL package shown in Figure 9 to include two QCLs, at different wavelengths and all the optical elements needed for dichroic beam combining (Figure 12).
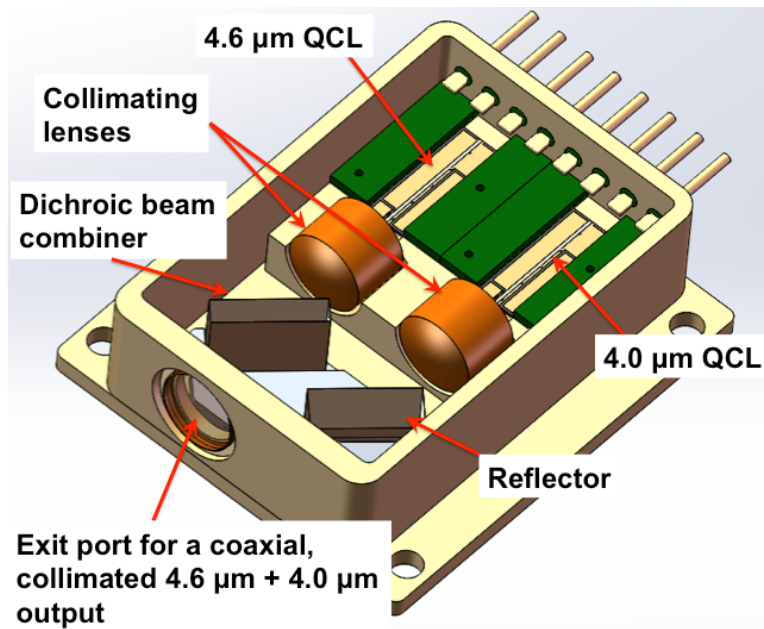

Figure 12. Compact package for beam combining of two QCLs at different wavelengths (Zebra package: 1.4" x 1.25" x 0.42")

We have successfully fabricated the compact package shown in Figure 11, with one QCL at 4.0 μm and the second QCL at 4.6 μm. The beam combining optical elements, the reflector and the dichroic mirror, are accurately placed and permanently affixed in place after alignment. Once aligned, there are no additional adjustments necessary or possible and the package is hermetically sealed to protect the QCLs and the optical elements from external environment. The package size is 1.4" x 1.25" x 0.42" and weighs ~45 g including all the optical elements. Each of the QCLs is energized using independent electronics drivers of the type shown in Figure 11, and thus the two QCLs

can be independently switched for simultaneous or otherwise operation. Table 4 shows experimental data for such a system.

Table 4. Performance data for a compact package using two different wavelength QCLs, the outputs of which are dichroically combined. Optical loss in beam combining is ~100 mW (~4%)

Laser	Wavelength	Power
# 1	~4.0 µm	1.0 W
# 2	~4.6 µm	1.5 W
Combined		2.4 W

Figure 13 shows a spectral scan of the combined beam output with both the QCLs operating simultaneously.
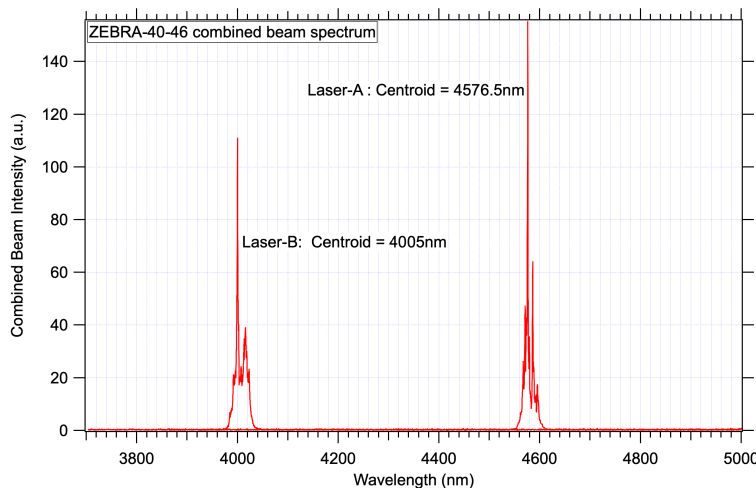

Figure 13. Spectral scan of output from the compact laser package shown in Figure 10.

Another application involves obtaining coherent power output at the same wavelength by beam combing outputs form two QCLs, to overcome the limitation of the single emitter power output. Of course, we cannot use dichroic beam combining, but we can take advantage of the fact that the laser output form QCLs is linearly polarized, thereby combine two orthogonally polarized beams from the two QCLs (at the same or different wavelengths) using a polarizer beam combiner. Again for laboratory applications, macroscopic optical elements, which are adjusted using optical mirror mounts, work quite well. However as with the Zebra package, such an arrangement is bulky, heavy and not rugged. To overcome these issues, as with the Zebra package, we have developed a compact coherent beam combining package shown in Figure 14.
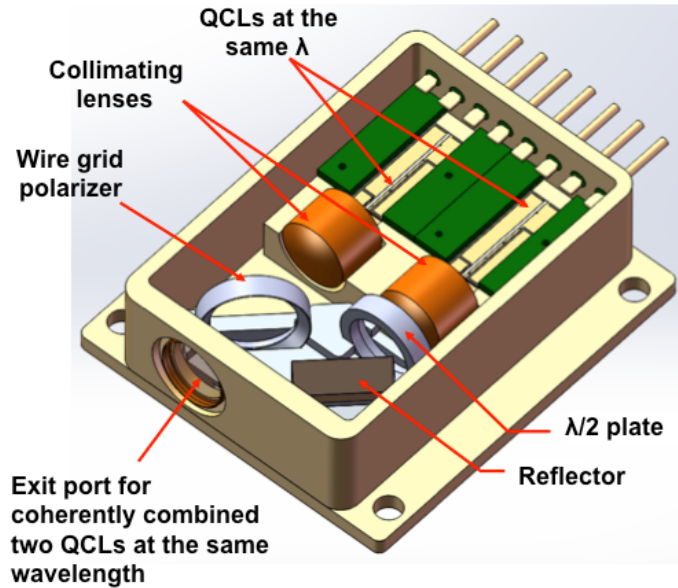
Figure 14. Compact package for beam combining of two QCLs at the same wavelengths (Dolphin package: 1.4" x 1.25" x 0.42")

In this case the two QCLs are at the same wavelengths and output from each of the QCL is vertically polarized. Output from one of the QCLs goes through a half-wave plate as shown, which rotates the polarization to horizontal. This beam is then reflected from the mirror and is incident on a wire grid polarizer that is adjusted to reflect horizontal polarization and transmit the vertically polarized light coming from the other QCL. The combined beam emerges from the package as a single coaxial beam but is unpolarized.   The power output is the sum total of the power output from the two QCLs, less losses in the optical elements. Table 4 shows performance of the arrangement, where the various optical elements permanently affixed after initial alignment and no further adjustment is either necessary or possible.   Subsequently the package is hermetically sealed. Table 5 shows performance of an early demonstration.

Table 5. Performance data for a compact package using two same wavelength QCLs, the outputs of which are polarization combined. Optical loss in beam combining is ~60 mW (~2%)

Laser	Wavelength	Power
# 1	~4.6 µm	1.73 W
# 2	~4.6 µm	1.39 W
Combined		3.06 W

Figure 15 shows a photograph of the combined beam output, at a distance of ~1.5 m from the exit of the package. Excellent beam overlap is evident and beam quality measurements on the combined single beam show a M^2 ~1.5.
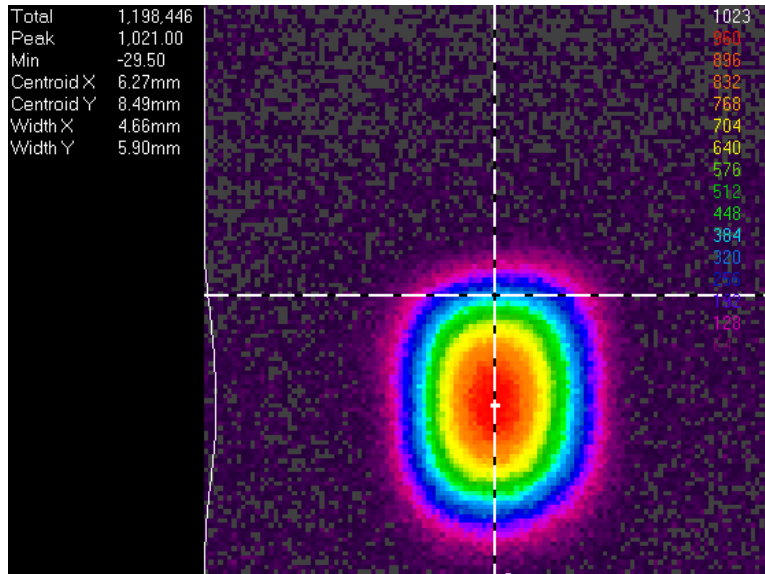
Figure 15. Photograph of combined beams at a distance of ~1.5 m from the exit from the multiple QCL package

With installation of two 2.5 W QCLs at 4.6 μm, we expect obtain nearly 5W of high beam quality output.

5. APPLICATIONS

QCLs are the only solid sources of laser radiation in the spectral region longer than ~3.5 μm that operate CW at room temperature that directly convert electrical power into optical power. The only other sources that convert electrical power directly into optical power in the ~3.5 μm to ~12.0 μm spectral regions are two principal gas lasers – the carbon dioxide laser [2, 3] that operates in the spectral region from ~9.0 μm to~11.5 μm, and the carbon monoxide laser [4] that operates in the spectral region from ~5.0 μm to~7.0 μm. Both of these lasers have been extensively developed and CO_2 lasers producing >10's kW of CW power and CO lasers producing >1 kW of CW power are commercially available. However, these lasers operate on vibrational-rotational transitions of respective molecules and thus the optical power output occurs only discretely spaced frequencies, which are ~4 cm^{-1} for CO and ~2 cm^{-1} for CO_2. Both of these lasers have found very large number of applications in defense and non-defense sectors. In non-defense sectors, these molecular gas lasers are work-horses in materials processing including welding, cutting, marking, and annealing. Because of the high powers available, these molecular lasers are extensively used in soft tissue surgery and related medical applications. It is clear that because the current power limitations, QCLs are not able to access most of these applications that require 10 W to 10's of kW of power.

However, the QCLs cover a much broader spectral region than the widely deployed CO_2 and CO lasers, and the QCL gain occurs over a wide spectral region making it possible to have continuously tunable (as opposed to step tunable) infrared radiation as we saw earlier in the discussion of tunable QCLs. Additionally, QCLs are being semiconductor laser have the advantage of small size, weight and power (SWaP) and ruggedness, which make them ideal for applications that may involve mechanical shocks and vibrations. Continuous tunability translates into advantage for spectroscopic environmental and industrial applications for monitoring trace concentrations of molecular gases in the MWIR and LWIR regions, since most of the important molecular species have their strong fundamental absorptions in these regions. Examples include monitoring of chemical warfare agents (CWAs), toxic industrial chemicals (TICs) and explosives vapors and remote (standoff) detection of improvised explosive devices (IEDs) in defense, and homeland security arenas, monitoring trace gas evolution during chemical reactions for process end point detection in chemical, biochemical and semiconductor industries and monitoring minor constituents of exhaled human gases for potential non-invasive disease detection.

The FP configuration QCLs produce higher power output than the tunable QCL systems, but in a much broader bandwidth and are non-tunable. However, the SWaP advantage, mentioned earlier, makes the compact, lightweight FP configured QCLs ideal for replacement of the currently used MWIR sources in a variety of applications, where until recently the needs were met by using optical parametric oscillators (OPOs) pumped by fiber lasers. These systems were bulky, heavy and power hungry, and precluded utilization in hand-held and battery operated applications. Bulk of these applications fall in the defense arena and include directed infrared countermeasures, target pointers and illuminators and IFF beacons.

Applications in defense and homeland security

Directional infrared countermeasures (DIRCM): With the proliferation of Man Portable Area Defense System (MANPADS), i.e., shoulder-fired ground-to-air missiles, danger from non-state actors acquiring these missiles for terrorism has been with us for some time. Until recently, the defense against such attacks on aircraft included deployment of chaff, etc. With the advent of MWIR laser sources, such as OPOs mentioned above, the defense became more active one by the aircraft based IR sources illuminating the seeks on the incoming MANPADS to lose its lock on the heat signature of the aircraft engine exhaust gases and make the MANPADS go away, eventually falling back to ground when the missile fuel supply is exhausted. As mentioned earlier, the SWaP advantage of QCLs, such as those shown in Figure 5 have become the natural laser sources for DIRCM. Defense organizations in US and elsewhere are now actively pursuing DIRCMs based on QCLs.

MWIR target pointers and markers: Optical marking of targets using a laser and subsequent elimination of the target by an agent who has the appropriate viewer to locate and target on the illuminated spot for accuracy provides enormous help to the

warfighters when operating in difficult environments. Until recently, such target pointers and designators used near infrared (NIR) laser sources and were successfully deployed. However, the viewer technology for "seeing" the NIR illumination has now become ubiquitous and all adversaries including state and non-state actors can easily obtained the NIR viewers (often called the night goggles), robbing our warfighters of the advantage of surprise when confronting an adversary. The advent of ultra lightweight and ultra compact packaged QCL systems (Figure 10) has made it possible to fabricate hand-held (or weapons mounted) battery operated MWIR pointers that operate in the first atmospheric window (relatively transparent spectral region for atmospheric transmission) from ~3.0 µm to ~5.0 µm. Viewers for these wavelengths are very limited and controlled in supply and thus the MWIR laser pointers have paved a way for our warfighters to reclaim the element of surprise, which was lost when using the NIR pointers.

MWIR IFF beacons: A corollary to the target pointers, described above, are Identify Friend or Foe (IFF) beacons, which are designed to prevent damage and casualty from friendly fire. Currently used IFF systems rely of NIR laser sources and as with pointers above, the proliferation of NIR viewers has resulted in loss of our advantage. Use of MWIR laser sources for IFF beacons can permit regaining the advantage. Furthermore, use of MWIR wavelengths is intrinsically more desirable because the atmospheric scattering loss varies as λ^{-4} and thus increases the viewing distance considerably, when compared with the range possible using NIR beacons.

Detection of CWAs and IEDs in battlefield environment: Passive hyperspectral analysis has been long used for the standoff detection of TICs and CWAs. Active standoff detection provides the next step in standoff detection, wherein a tunable laser illuminates an area and the detection is accomplished by evaluating the return signal. Tunable lasers are ideal sources for active standoff detection. Line tunable carbon dioxide lasers, because of their relatively high power, have been extensively used for standoff detection starting with the detection of trace amounts of explosives (TNT) up to a distance of ~150 m and detection of trace amount of gaseous contaminant (SF_6) up to distance of 400 m [14]. Subsequently a CO_2 laser based active sensor for the detection of CWAs has also been developed. However, as mentioned earlier, a line tunable CO_2 laser provides radiation only from ~9.0 µm to ~11.5 µm. Availability of tunable QCL has now changed the picture significantly.

The broad spectral tunability of QCLs described earlier has been exploited for *in situ* detection of CWAs at very low levels for assuring the safety of military personnel as well as civilians. Detection of several CWA surrogates and TICs at ppb level has been demonstrated [15, 16].   Imaging stand-off detection of explosive substances on various surface at distances of the order of meters has also been carried out [17].

Detection of IEDs, especially in a battlefield environment poses a special problem. For example, a military vehicle (a Humvee) travelling at 30 mph needs to examine a roadside object to determine whether it is an IED or not very quickly, in order for the

operator to stop the vehicle. Furthermore, any sensor mounted on the Humvee needs to be able to withstand mechanical vibrations and shocks and operate while the vehicle is moving. As mentioned earlier, use of the AOM tuned QCL system (VeloXscan, Figure 9) provides the required data in ~100 µs, during which the Humvee will have moved less than inch, providing sufficient time for the operator to maneuver appropriately.

Applications in non-defense sectors

Process end-point detection: In chemical, biochemical and semiconductor industries many of the processes during production and processing are accompanied by evolution of trace gases, determination of whose presence, absence or quantitative concentration, could provide very valuable information about the status of the process itself. Instead of relying on time or past experience based determination of the process time, monitoring the concentrations of key gaseous components during the process should provide real time information about the status, which will provide a pathway to improved processes and better yields, both of which are critical to efficient production and processing. An example of such real time process control could be monitoring carbon dioxide evolution during fermentation of grapes (wine making). As long as the sugar is being converted into alcohol, the yeast mediated fermentation process results in carbon dioxide production. For a desired level of residual sugar, the wine maker normally draws a sample of liquid out of the fermenter to check for alcohol and sugar content. This process can be easily automated by real time measurement of the CO_2 concentration in the evolved gases. Similarly, in silicon semiconductor integrated circuit manufacturing, there are many processes during which there is etching and chemical deposition of materials. These processes can also be monitored in real time for improving the yields of devices.

Non-invasive medical diagnostics: It has been long recognized that exhaled human breath contains as many a few hundred trace gases that escape through the blood-lung barrier and some of these are indicators of health of the individual. Based on this observation, there are at least two specialized human breath gas sensors available that detect the presence of nitric oxide and ammonia in human breath. Elevated nitric oxide presence is an indicator of asthmatic attack while elevated ammonia is an indicator of high concentration of creatinine and ammonia in blood signifying kidney and/or liver dysfunction. The NO sensor uses a chemical titration technique for ppb level detection, while the ammonia sensor uses a CO_2 laser based spectroscopic technique. With the availability of broadly tunable QCL systems, it is not too far fetched to expect a non-invasive breath diagnostic system that monitors a larger number of trace gases simultaneously providing a broader picture of the patient's health. Of course, detection of intoxicating substances in human breath, such as those present in individual consuming alcohol or inhaling cannabis smoke, should be of interest to law enforcement agencies. Detection of both of these is well within the capability of a QCL based spectrometer.

Agriculture: Passive hyperspectral analytic tools have been used for remote monitoring soil moisture extensively. As mentioned earlier, use of laser based active standoff detection will improve the standoff detection capability significantly by being able to measure only the soil moisture but also other low level emanations from the agricultural crops that may be indicative of crop disease, infestation and/or the crop status, especially for ripening of fruits in a large farm. With the availability of the ultra light weight and ultra small high power QCLs, one should be able to install a number of these on a small UAV to capture the required spectral information for real time monitoring the farm status.

Food industry: Food production and distribution industry involves very storage houses that are appropriately environment controlled. Nonetheless, there are large number of examples of E. Coli contamination and spoilage of stored food stuff in large stage houses. All degradation in the quality of food is accompanied by release of signature gases that can be easily monitored using QCL based spectroscopic techniques, providing the industry advance notice of problems that need to be addressed before the spoilage leads to economic loss as well as health crises.

Free space optical communications: Because of the intrinsic QCL laser level lifetimes being of the order of picosecond or less (see earlier discussion of QCL structure design), QCLs can be directly modulated at rates of tens of gigahertz. This capability can be exploited for high data rate free space optical (FSO) communications using the MWIR and LWIR quantum cascade lasers. Early studies have already confirmed the advantage of MWIR/LWIR FSO communications in its ability to operate in highly scattering environmental conditions [18, 19]. While FSO communications using QCLs haven't yet reached the commercial or defense markets, it is very possible that some of the improvements that will be described later in this paper about future path, could make QCL based FSO communications economically attractive.

6. FUTURE

As can be seen, QCLs have come a long way in the 25 years since the first demonstration. QCLs have made significant inroads into a broad range of applications, in defense as well as non-defense areas. However, compared with the extremely broad deployment of carbon dioxide lasers (the other room temperature infrared laser that directly converts electrical power into optical power), within very few years from its invention, QCL applications lag significantly behind. This is in spite of the many of the advantages QCLs hold over other lasers in the MWIR/LWIR spectral region. We see that there are at least three factors where we need additional focus. These are:

- **Performance improvement:** We are, at present stuck at about 5 W per single emitter (hero devices) and ~3.0 W for production devices. Structure design, processing and packaging innovations are needed to increase the single emitter power levels to at least 10 W for production devices. At this power level, QCLs have a significant opportunity to displace many of the low power CO_2 lasers because of

the SWaP advantage (in addition to the ruggedness that is characteristic of semiconductor lasers). The increase in single emitter power output will require rethinking fundamental QCL structure design for higher wall plug efficiency, since thermal management is a very critical part of increasing single emitter power. One of the important steps that is being explored at present is the quantify the improvement that is possible if we use 10 to 30 gain stages in a QCL as opposed to the current 40-50 gain stages [8].

- **Reliability, reproducibility and lifetime improvement:** Reliability and lifetime, at present are acceptable for systems that would go through upgrade routinely, such as those in the defense arena. However, for non-defense applications, we need a substantial improvement to bring the high power QCL lifetimes to the near IR semiconductor lifetimes of >40,000 hours. Reproducibility of performance is another key parameter that need attention. No matter how reliable a device may be, there will always be a failure in a deployed system. A system operator would like to replace the malfunctioning device with a new one, without having to "retune" the system parameters. For QCLs, this translates into all QCLs, for a particular application, having identical LIV characteristics, threshold current, maximum power and power at operating current (which is set in the system).

- **Cost:** We, as many others, have identified the root cause of relatively slow growth of applications of QCLs. This is the cost of a fully packaged and tested QCL. The present cost is ~$20,000.00 per watt for a high performance QCL. And the cost scales faster than linearly with increasing power. Thus, at present the most significant applications are where the payoff, either in prevention of damage or in improvement in the processes being controlled, is such that cost of the QCL is not a major factor. This is especially true for many defense applications where performance trumps the cost. However, for many other defense and non-defense applications cost is an equal partner with performance. We believe that bringing down the cost of a fully packaged QCL to less than $10,000.00 per watt (from the present $20,000.00 per watt) would increase the market opportunity by more than factor of ten. The cost of the QCL, as with any other device, is closely connected with the yield of QCLs in the production process.

Of course power is not the only parameter that determines the yield and therefore the cost of the QCL. In addition to power level, two additional factor affect the yield and these are spectral content of the output and the output beam quality. Table 6 shows how yields (therefore the cost) are affected by the three key performance parameters of QCLs.

Table 6. Yield as a function of performance parameters

REQUIREMENT	YIELD				
	0.5 W	1.0 W	2.0 W	3.0 W	4.0 W
Power level only	50%	40%	10%	3%	<1%
Power level+spectral content	40%	20%	2%	<1%	N/A
Power level+spectral content+beam quality	20%	10%	1%	<1%	N/A

We have identified twelve major steps (Figure 16) critical for yield from QCL structure design to final fully packaged QCL.
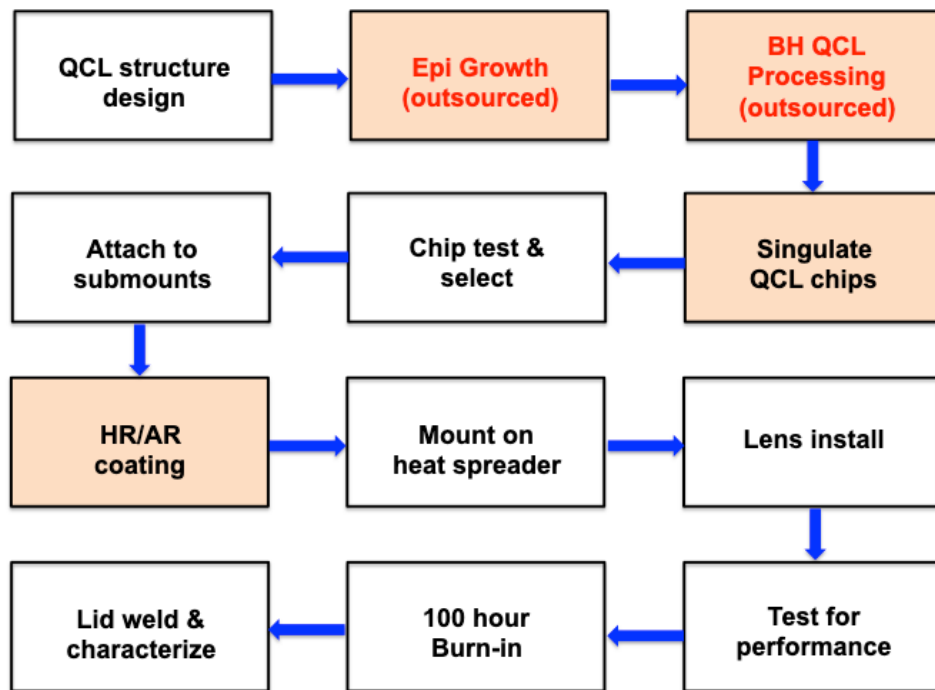

Figure 16. QCL production flow

Steps, which result in significant yield loss, are shown in red hue in Figure 16:

- MBE growth of epi wafer: ~2000-2500 n-type layers of $In_x Ga_{1-x} As$ and $In_y Al_{1-y} As$ with precise thickness and doping control (problems of lack reproducibility of design wavelength, power, lasing threshold, etc.) for 40-50mgain stage structure design
- Buried heterostructure (BH) laser processing (very deep etching and regrowth of Fe:InP)
- Singulating QCL chips from a 2" or a 3" processed wafer and attaching to submounts
- HR/AR coating on end-facets (primary cause of catastrophic failure of QCLs at high powers)

QCL production, because of the current low volume requirements, has not addressed the yield management issue and there is very little documentation of "cause and effect" that can be fed back into the process for yield improvement. To address the issues, we provide the following step-by-step guidance that is reminiscent of what happened in silicon IC industry for yield improvement.

–   Measure key parameters in each of the processing steps, especially in those, which are identified as the critical ones.
–   Define yield at each step
–   Correlate yield with process parameter variations
–   Carry out the required parameter determinations and yields at each step
–   Create a statistically relevant Yield Management Algorithm

Substantial benefits will accrue to the whole QCL community, once we decide that yield improvement will lead to better QCLs at a lower cost.

REFERENCES

[1] J. Faist, F. Capasso, D. Sivco, C. Sirtori, A. Hutchinson, and A. Cho, "Quantum cascade laser," *Science*, vol. 264, pp. 553–556, 1994.

[2] C. K. N. Patel, "Interpretation of CO_2 Optical Maser Experiments", *Phys. Rev. Lett.* **12,** 588-590 (1964).

[3] C. K. N. Patel, "Continuous-Wave Laser Oscillation on Vibrational-Rotational Transitions of CO_2" *Phys. Rev.* **136A,** 1187-1193 (1964).

[4] C. K. N. Patel and R. J. Kerl, "Laser Oscillation on Σ^1S^+ Vibrational-Rotational Transitions of CO", *Appl. Phys. Lett.* **5,** 81-83 (1964).

[5] Mattias Beck, Daniel Hofstetter, Thierry Aellen, Jérôme Faist, Ursula Oesterle, Marc Ilegems, Emilio Gini, Hans Melchior, "Continuous Wave Operation of a Mid-Infrared Semiconductor Laser at Room Temperature", *Science* **295**, 301-305 (2002)

[6] Arkadiy Lyakh, Richard Maulini, Alexei Tsekoun, Rowel Go and C. Kumar N. Patel, "Tapered 4.7 µm quantum cascade lasers with highly strained active region composition delivering over 4.5 watts of continuous wave optical power" *Optics Express* **20 (4)** 4382-4387 (2012).

[7] C. Kumar N. Patel, Arkadiy Lyakh, Alexei Tsekoun and Richard Maulini, "Submounts for Semiconductor Lasers" *U. S. Patent* # 8,068,524 (Nov. 29, 2011).

[8] Pedro Figueiredo, Matthew Suttinger, Rowel Go, Ankesh Todi, Hong Shu, Eugene Tsvid, C. Kumar N. Patel and Arkadiy Lyakh, "Continuous Wave Quantum Cascade Lasers with Reduced Number of Stages", *IEEE Photonics Tech Lett.* **22 (16),** 1328-1331 (2017)**.**

[9] Arkadiy Lyakh, Matthew Suttinger and Rowel Go, "High-brightness, broad area continuous wave quantum cascade Lasers" Proc. *SPIE* (Presentation 11161-22) 2019.

[10] Andreas Hugi, Gustavo Villares, Stephane Blaser, H. C. Liu and Jérôme Faist, "Mid-infrared frequency comb based on a quantum cascade laser", *Nature* **492,** 229-233 (2012).

[11] Jérôme Faist, Gustavo Villares, Giacomo Scalari, Markus Rösch, Christopher Bonzon, Andreas Hugi, and Mattias Beck, "Quantum Cascade Laser Frequency Combs" *Nanophotonics* **5,** 271-291 (2016).

[12] Arkadiy Lyakh, Rodolfo Barron-Jimenez, Ilya Dunayevskiy, Rowel Go and C. Kumar N. Patel, "External cavity quantum cascade lasers with ultra rapid acousto-optic tuning" *Appl. Phys. Lett.* **106**, 141101 (2015).

[13] C. Kumar N. Patel, Rodolfo Barron-Jimenez, Ilya Dunayevskiy, Gene Tsvid and Arkadiy Lyakh, "Two Wavelength Operation of an Acousto-Optically Tuned Quantum Cascade Laser and Direct Measurements of Quantum Cascade Laser Level Lifetimes", *Appl. Phys. Lett.* **110,** 0331104 (2017).

[14] Anadi Mukherjee, Steven Von der Porten and C. Kumar N. Patel, "Standoff detection of explosive substances at distances of up to 150m", *Appl. Opt.* **49,** 2072-2078 (2010).

[15] Anadi Mukherjee, Manu Prasanna, Rowel Go, Ilya Dunayevskiy, Alexei Tsekoun and C. Kumar N. Patel, "Optically Multiplexed Multi-gas Detection using Quantum

Cascade Laser Photoacoustic Spectroscopy", *Applied Optics* **47,** 4884-4887 (2008).

[16] Coorg R. Prasad, Jie Leia, Wenhui Shi, Guangkun Lib, Ilya Dunayevskiy and C. Kumar N. Patel, "Laser Photoacoustic Sensor for Air Toxicity Measurements", *Proc SPIE* **8366,** 836608 (2012).

[17] Frank Fuchs, Stefan Hugger, Michel Kinzer, Rolf Aidam, Wolfgang Bronner, Rainer Losch and Quankui Yang, "Imaging standoff detection of explosives using widely tunable midinfrared quantum cascade lasers", *Optical Engineering* **49(11),** 111127 (2010).

[18] R. Martini, C. Gmachl, J. Falciglia, F.G. Curti, C.G. Bethea, F. Capasso, E.A. Whittaker, R. Paiella, A. Tredicucci, A.L. Hutchinson, D.L. Sivco, and A.Y. Cho, "High speed modulation and free space optical audio/video transmission using Quantum Cascade lasers", *Electronics Letters* **37 (3),** 191-193 (2001).

[19] R. Martini and E.A. Whittaker, "Quantum cascade laser-based free space optical communications" *J. Opt. Fiber. Commun. Rep.* **2,** 279–292 (2005).