

# A new high-efficiency regime for gas phase terahertz lasers: Experiment and ab-initio theory

Fan Wang,<sup>1,2</sup> Jeongwon Lee,<sup>2</sup> Dane J. Phillips,<sup>5</sup> Samuel G. Holliday,<sup>5</sup> Song-Liang Chua,<sup>6</sup> Jorge Bravo-Abad,<sup>7</sup> John D. Joannopoulos,<sup>2,3</sup> Marin Soljačić,<sup>2,3</sup> Steven G. Johnson,<sup>2,3,4</sup> Henry O. Everitt<sup>8,9</sup>

<sup>1</sup>*Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

<sup>2</sup>*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

<sup>3</sup>*Department of Physics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*

<sup>4</sup>*Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*

<sup>5</sup>*IERUS Technologies, 2904 Westcorp Blvd, Ste 210, Huntsville, AL 35805, USA*

<sup>6</sup>*DSO National Laboratories, 20 Science Park Drive, Singapore 118230, Singapore*

<sup>7</sup>*Departamento de Física Teórica de la Materia Condensada and Condensed Matter Physics Center, Universidad Autónoma de Madrid, 28049 Madrid, Spain*

<sup>8</sup>*Department of Physics, Duke University, Durham, NC 27708, USA*

<sup>9</sup>*Army Aviation & Missile RD&E Center, Redstone Arsenal, Alabama 35898, USA*

**We present both a new theoretical description and an experimental validation of a molecular gas optically pumped far infrared (OPFIR) laser (Fig. 1) at 0.25THz that exhibits 10×**

**greater efficiency (39% of the Manley–Rowe limit) and 1000× smaller volume than comparable commercial lasers (Fig. 2). Unlike previous OPFIR-laser models involving only a few energy levels<sup>1–4</sup> that failed even qualitatively to match experiments at high pressures, our ab-initio theory matches experiments quantitatively with no free parameters (Fig. 3) by accurately capturing the interplay of millions of degrees of freedom in the laser. We show that previous OPFIR lasers were inefficient simply by being too large, and that high powers favor high pressures and small cavities. We believe that these results open the possibility of a new generation of compact THz sources.**

For many years, OPFIR gas lasers were the most powerful sources of continuous-wave (CW) terahertz (THz) radiation,<sup>1–13</sup> and such THz sources are crucial to a wide variety of sensing and imaging applications.<sup>14–16</sup> They employ population inversions excited in molecule-specific rotational–vibrational transitions by a line-tunable CO<sub>2</sub> laser. By changing the gas, isotope, and CO<sub>2</sub> laser line, different transition frequencies could be made to lase, typically in meter-long cavities operated at low pressures (< 0.1 Torr) and producing up to 100 mW in a manner that depended strongly on the rotational transition involved.<sup>17,18</sup> Initially, it was thought that low pressures and large cavities were required, but this produced a “vibrational bottleneck” arising from diffusion-limited relaxation by molecule–wall collisions that quenched the inversion as pressure increased.<sup>2,4,7</sup> It was subsequently discovered that OPFIR lasers could operate at higher pressures in smaller cavities as a result of collisional processes that excited high-energy vibrational levels.<sup>8,10</sup> This breakthrough was never fully investigated or exploited, however, as it came at a time when new sources of terahertz radiation were emerging<sup>19–21</sup> and the use of OPFIR lasers was waning due to their supposed

large sizes (e.g. 10cm in diameter and 1m in length) and low photon conversion efficiencies (e.g. 0.1-10%).

The advent of alternatives to line-tunable CO<sub>2</sub> lasers<sup>22,23</sup> and the increasing need for powerful, narrowband sources of terahertz radiation is reviving interest in compact OPFIR lasers,<sup>1</sup> whose full potential has never been fully understood because of the thousands of quantized states involved and the millions of collision cross sections required to simulate its behavior adequately. Here we solve this problem, providing an accurate theoretical model with no adjustable parameters that quantitatively matches new experimental measurements of a compact OPFIR laser. Not only does the model match experiments, but it also unexpectedly shows that our compact laser exhibits remarkable efficiency at high pressures, more than 10× what is possible with the best commercial OPFIR lasers (Fig. 2).

Figure 1 depicts our OPFIR laser: a line-tunable CO<sub>2</sub> pump laser ( $\lambda = 9.7\mu\text{m}$ ) is focused through a pinhole in the front copper mirror into a 5mm×14.3cm copper tube filled with <sup>13</sup>CH<sub>3</sub>F gas. The lasing transition is brought into resonance with the cavity by a copper-plunger rear mirror that adjusts the cavity length, and the resulting laser power emerges from the same pinhole, after which it is detected by a heterodyne receiver. The different rotational levels are characterized by a rotational quantum number  $J$ , its projection quantum number  $K$ , and the vibrational state  $V_\ell$  to which it belongs.<sup>24,25</sup> For symmetric-top molecules, there are two degenerate symmetry types, A and E,<sup>11,24</sup> due to nuclear-spin interactions. In <sup>13</sup>CH<sub>3</sub>F, the pump process only couples to type-A molecules: the coincidence between the 9P(32) line of the pump laser and the R<sub>3</sub>(4) ro-vibrational

transition in  $^{13}\text{CH}_3\text{F}$  produces two rotational population inversions, the “direct”  $J = 5 \rightarrow 4$  (with  $K = 3$ , type A) inversion in  $V_3$  with frequency 245.351 GHz, and a corresponding “refilling” rotational transition in the type A ground vibrational level  $V_0$  with frequency 248.559 GHz (Fig. 1).<sup>1</sup> The heterodyne receiver distinguishes the lasing transitions and measures their intensity as a function of pressure and the input pump power. Figure 3 shows that the refilling transition produces the most terahertz power in a manner that depends nonlinearly on the input pump power, but the direct transition thrives over a much broader range of pressures. In contrast with traditional OPFIR lasers, these pressure-dependent behaviors in our compact OPFIR laser do not bottleneck with increasing pump power on either transition, and increasing power allows higher-pressure operation, so even stronger emission and higher pressures are achievable.

To understand the lack of power/pressure saturation, a thorough model of OPFIR laser operation must capture many different physical processes. Gas-phase molecules typically have hundreds to thousands of thermally populated rotational energy levels in the same vibrational modes as the pumped states ( $L \rightarrow U$ ). It has previously been shown that rapid collisional thermalization maintains a Boltzmann distribution of population across most rotational states within a given vibrational level  $V_\ell$  and symmetry type (A or E); consequently, they may be modeled as a thermal pool.<sup>6,9–11</sup> This thermal-pool approximation dramatically simplifies models of the collision physics by reducing the number of levels that must be individually included and collision cross sections that must be known to those few most closely connected to L and U. Nevertheless, after the pump excites molecules from L to U, a complicated set of collisional relaxation processes ensue, both bimolecular rotational energy transfer and molecular diffusion to the walls for ro-vibrational thermalization,

so that population inversions between  $L+1$  and  $L$  (refilling) and  $U$  and  $U-1$  (direct) occur in a manner that depends sensitively on the input pump power, pressure, radial population distribution, cavity length, and cavity mode. Because wall collisions are the primary pathway for vibrational relaxation in OPFIR lasers and the diffusion rate goes as  $1/pR^2$  for pressure  $p$  and cell radius  $R$ , typical large-diameter glass OPFIR lasers require operation at low pressure, while small-diameter metallic cells are favored for high-pressure operation with the tradeoff of increased ohmic losses at the lasing frequency.

Various models have been developed in an attempt to understand the inversion process.<sup>1-4</sup> A troubling limitation in all these models, shown in Fig. S2 in the Supplementary Information (SI), was that they exhibited an artificial bottlenecking of the gain at high pressures, *regardless* of the input pump power, rendering them unable to explain high-pressure lasing such as that shown in Fig. 3. Ref. 8 addressed the problem of high-pressure bottlenecking by including more vibrational levels, but the model was too oversimplified and incomplete to provide useful quantitative guidance for cavity design, and the lack of an efficient numerical solver prohibited any attempt at a more realistic model.

Retaining the spirit of Ref. 8, we construct a far more complete model that quantitatively describes how this artificial bottlenecking is overcome through the collisional excitation of molecules into hundreds of excited vibrational levels in compact OPFIR lasers. In our new model, whose complete mathematical details are provided in Methods, the degrees of freedom in the rate equations include axial molecular speed  $v$ , radial position  $r$ , time  $t$ , and different energy levels including

nonthermal rotational levels and vibrational thermal pools. By calculating separate populations at each  $v$ , unlike previous work, we can accurately model inhomogeneous broadening and spectral hole burning (SHB) effects on the Doppler-broadened pumped infrared transition and the resulting pressure-broadened terahertz gain profile. By additionally modeling each population density as a function of  $r$ , the overlap between the inversion and the cavity mode is calculated for an accurate estimation of the weighted gain coefficient. In the efficient regime where the pump intensity reaches the entire cavity (i.e. infrared absorption, spectral hole burning, and saturation effects do not cause the pump intensity to decay strongly before reaching the end of the cavity), the populations can be approximated as uniform in  $z$  with an effective  $z$ -averaged pump intensity (different from the input pump intensity, see SI), which greatly simplifies the calculations. Given a set of nonlinear rate equations, most authors evolve them in time until the steady state is obtained,<sup>1,10</sup> but this approach is severely inefficient, especially at high pressures where timescales for different relaxation processes diverge widely. Instead, we solve for the steady state directly, without time evolution, as a system of millions of coupled nonlinear equations, with computational cost equivalent to less than 100 time-evolution steps. The enormous computational improvement allows us to include not only millions of population degrees of freedom but also many physical processes that other authors neglected or approximated. In particular, we include dipole–dipole collisions between rotational states,<sup>13</sup> symmetry-preserving thermalization (SPT, also called K-swap processes) between nonthermal rotational states and vibrational thermal pools, V-swap process,<sup>11,24</sup> non-resonant collisions among thermal pools,<sup>24,25</sup> molecule–wall collisions,<sup>1,10</sup> and spatial diffusion.<sup>1</sup> For the inhomogeneously-broadened pump absorption,<sup>25</sup> we include Doppler, AC-Stark,<sup>26</sup>

and pressure<sup>27</sup> broadening of velocity subclasses, and we determine the average pump intensity self-consistently, including SHB and pump saturation, with the absorption coefficient calculated from the populations. Given rate equations for the populations  $N_k(v, r, t)$  of each level  $k$  (see Methods), by setting  $\partial N_k/\partial t = 0$  we obtain a nonlinear system of equations for the steady-state populations  $N_k(v, r)$ . From the calculated radially-dependent population inversions, gain profiles, and lasing mode geometry, the output of the model is the predicted lasing power  $I_o$ , calculated using modified standard formulas<sup>27</sup> with a novel self-consistent characteristic time as described in Methods.

Even with the thermal pool simplification and steady-state computation efficiency, the sheer complexity of the inversion process makes it extremely challenging to model accurately. Previous models needed to make many simplifying assumptions, and the reason they produced artificial bottlenecking at high pressures was conjectured to be their inclusion of only two key vibrational levels,<sup>8,10</sup> e.g.  $V_0$  and  $V_3$  for  $^{13}\text{CH}_3\text{F}$ . The difficulty of including more vibrational levels seems almost insurmountable at first: although the energies of hundreds vibrational levels are known,<sup>25</sup> the relaxation processes among them are not known quantitatively and are infeasible to measure. To deal with that problem, a crucial theoretical innovation of our work is to describe those levels implicitly via an effective temperature (different from the ambient temperature!) that is determined self-consistently from the few collisional cross sections whose values are known. Specifically, we make the ansatz that vibrational transitions are rapid enough to “thermalize” those levels with a Boltzmann distribution described by an effective temperature  $T_v$ , so they may be grouped together into a population pool  $V_\Sigma$ , for both A and E symmetries, as shown in Fig. 1(b). This is similar to the

experimentally justified thermal-pool assumption<sup>9,11</sup> that maintains a constant temperature among rotational states within a given symmetry type and vibrational level. The main justification of this ansatz is that thermalization among vibrational states is collisional, meaning the rates increase with pressure, while vibrational-state relaxation occurs through diffusion to the walls, which goes as the inverse of pressure. Therefore, the higher the pressure, the more vibrational thermalizing collisions occur before diffusion to the walls relaxes the excitation, and so our model becomes more accurate precisely in the high-pressure regime where these extra vibrational levels are relevant to the inversion.

To implement this thermalization ansatz, the populations of  $V_0$  and  $V_3$  are calculated explicitly, and we then assign the effective temperature of the remaining vibrational levels to be  $T_v = (E_{V_3} - E_{V_0})/k_B \log(N_{V_3}/N_{V_0})$ , where  $E_{V_\ell}$  and  $N_{V_\ell}$  are the energy and population of level  $V_\ell$ .  $T_v$  must be determined self-consistently by our nonlinear solver, simultaneously with  $N_{V_3}$  and  $N_{V_0}$ , because any population assigned to  $N_{V_\Sigma}$  according to  $T_v$  is removed from  $N_{V_3}$  and  $N_{V_0}$ . If we replace these effective vibrational levels with an explicit 6-level model (or even a 3-level model as in previous work<sup>1-4</sup>), but include all the other processes of our model, we find that the gain artificially bottlenecks similar to previous work: the gain disappears above 700 mTorr for all input pump powers (see Fig. S2 in SI). Therefore, we confirm the conjecture that including a complete set of vibrational levels through  $V_\Sigma$  eliminates high-pressure bottlenecks. Indeed, Fig. 3 shows excellent agreement between the experimentally measured power (a) and the predictions of our model (b), with no free parameters except for the overall scaling, even for pressures  $> 1$  Torr where previous models exhibited bottlenecks. In the dotted parts of Fig. 3(b), the pump is highly focused in the

front of the cavity and quickly decays to nearly zero ( $< 2\%$ ) by the back of the cavity, in which case our model’s approximation of an effective  $z$ -uniform pump strength is inaccurate, and hence the slight mismatch with experiment is unsurprising. Accurately modeling this attenuated-pump regime is relatively unimportant because, as described below, that regime “wastes” the end of the cavity, and it is preferable to use a shorter cavity.

The agreement between theory and experiment allows us to use the theory to assess the attainable quantum efficiency of the laser accurately and quantitatively, a parameter which was difficult to measure in our experiment because only a fraction of the emitted photons were detected by our heterodyne receiver (see Methods). Figure 3(c) plots the predicted lasing threshold and differential quantum efficiency (DQE) at the threshold, defined as  $\eta_d = \frac{dP_{\text{out}}}{dP_{\text{pump}}} / \frac{\nu_{\text{out}}}{\nu_{\text{pump}}}$ . The figure shows that the direct transition has lower pump threshold, but its DQE is smaller than the refilling transition (which can reach as high as 64%). We also obtain the total quantum efficiency (QE), defined as  $\eta_t = \frac{P_{\text{out}}}{P_{\text{pump}}} / \frac{\nu_{\text{out}}}{\nu_{\text{pump}}}$ , which is theoretically bounded by 100% from the Manley–Rowe limit:<sup>27</sup> the highest achievable power conversion efficiency is the ratio of output and pump frequencies. Our computed QE can be as high as 29% at 350 mTorr for the refilling transition and, as explained below, can increase to 39% in an optimized cavity. As shown in Fig. 2, both QE values are considerably larger than those attained in previous commercial OPFIR lasers.

The remarkably high QE and DQE are an unexpected benefit of high pressure operation, especially for the refilling transition. The model indicates that as pressure increases,  $T_v$  also increases as more of the pumped molecules are collisionally transferred from  $V_3$  to  $V_\Sigma$  and are

therefore removed from the  $V_0$  and  $V_3$  A-type thermal pools that quench the refilling and direct inversions, respectively. DQE measures the likelihood that each additional pump photon above threshold produces an additional terahertz photon at the threshold, and the high DQEs seen at the highest operational pressures mean that both inversions are increasingly efficient as bottlenecking traps molecules in higher vibrational levels. In other words, below threshold most pumped molecules are quickly lost from the levels associated with the inversions due to rapid collisional thermalization, but just above threshold for a given pressure, sufficient collisional redistribution and equilibration of population has been achieved that additional pump photons simply add to the strength of the inversions. This is particularly true for the refilling transition, where the inversion with  $L+1$  is produced as the pump laser removes molecules from  $L$ . The efficiency of the direct transition is lower and less sensitive to pressure because the pumped molecules placed in  $U$  are quickly removed by collisional processes. Consequently, the model indicates that it is the refilling transitions, not the direct transitions typically used in traditional OPFIR lasers, that are most attractive for compact, high pressure, efficient OPFIR lasers.

With this full understanding of the lasing physics at all pressures, we can finally explore and optimize the cavity geometry to maximize the THz power for compact OPFIR lasers. In particular, we consider the choice of cavity length and radius. Performing parameter variation with our model indicates that output power is maximized by a cavity length of roughly  $1/2\alpha_p$ , where  $\alpha_p$  is the average pump absorption coefficient (see SI), so here we define  $L_o = 1/2\alpha_p$  to be the “optimal” cavity length. Fortunately, operating with this cavity length also ensures the validity of our  $z$ -averaged pump-intensity approximation, since  $L_o$  is several times smaller than

the length where this approximation begins to fail due to nonuniform pump intensity. Calculating  $L_o$  is nontrivial because  $\alpha_p$  itself depends on cavity length (see SI), so  $L_o$  must be obtained by a self-consistent nonlinear search. Figure 4(a) shows the optimal cavity length as a function of pressure with three different radii  $R = 0.15\text{cm}$ ,  $0.25\text{cm}$ , and  $0.4\text{cm}$ , with pump power set at  $10\text{W}$ . Because pump absorption increases with increasing pressure, we can see that higher pressures favor shorter cavities. The increase in optimal cavity length with decreasing cavity diameter derives from the increasing pump intensity, the associated increase in AC-Stark broadening and SHB, and the increasing saturation of the pumped  $L \rightarrow U$  transition for a given input pump power. Given that our cavity length was maintained at approximately  $14.3\text{cm}$  for all our measurements, the model indicates that the experimental drop in OPFIR power with increasing pressure is caused in part by excess cavity length and incomplete pumping of the entire volume. Another design improvement can be achieved by placing the front pinhole at an optimized position: one chooses the pinhole location to increase the output flux of the lasing mode, which increases both the output power and the efficiency by decreasing the fraction of power lost to Ohmic absorption. For example, the  $\text{TE}_{01}$  mode<sup>28</sup> has a maximum flux intensity around  $r = 0.48R$ , so an off-center pinhole will achieve larger lasing power, as shown schematically in Fig. 1(a). We compute the effect of the pinhole location by calculating the  $\text{TE}_{01}$  flux through the hole in SI.

With the input pump power set at  $10\text{W}$ , the predicted output power for cavities of optimized length and pinhole position is shown in Fig. 4(b). Comparison with Figs. 3(a,b) indicates how severely the high-pressure lasing is hampered by excess cavity length. For optimized cavities, the maximum output power can reach  $31\text{mW}$ , which is  $39\%$  of the Manley–Rowe limit. Cavities with

a larger radius of  $R = 0.4\text{cm}$  can achieve slightly higher output power at low pressures (around 100 mTorr). However, low pressure indicates longer cavity length (around 50cm) from Fig. 4(a). In contrast, 0.15cm diameter cavities shorter than 20cm can generate output power about 25% of the Manley–Rowe limit in a wide range of pressure above 400 mTorr. This requirement for widely adjustable cavity lengths is in stark contrast to traditional fixed geometry OPFIR lasers, typically 10cm in diameter and 1 meter in length.<sup>17</sup> Our analysis explains why these huge cavities are inefficient, can only operate at very low pressures, and may not be long enough!

Our compact laser is  $1000\times$  smaller in volume while delivering  $10\times$  greater efficiency than previous OPFIR lasers at this wavelength. This is shown in Fig. 2, where the quantum efficiency normalized by the Manley–Rowe limit is plotted for both our experimental setup and our optimized compact laser, as well as for the best commercial OPFIR lasers. The Manley–Rowe limit indicates that for a given input pump power and quantum efficiency, output power increases linearly with decreasing lasing wavelength, so our analyses suggest that future OPFIR lasers can produce even higher powers with a careful choice of cavity and gas. Moreover, we have only scratched the surface of the design possibilities offered by our ability to model the full physics of OPFIR lasers accurately, and we believe that many further discoveries await the extension of our approach to new gases (e.g. for shorter-wavelength operation) and new cavity designs.

## Methods

**Experimental measurements.** The pump beam from a grating-tuned Apollo CO<sub>2</sub> laser (Model 570), producing as much as 15W on the 9P(32) line at  $9.657\ \mu\text{m}$ , was focused through a 1mm

diameter pinhole in a gold-coated copper plate by a 10cm focal length ZnSe lens into the 0.5cm diameter, 14.3cm long OPFIR laser cavity with a base pressure of  $< 10^{-4}$  Torr and a transmitted laser power 2/3 of the incident power (e.g. 10W input pump power at maximum output). The cavity was composed of copper tubing with a gold-coated cylindrical copper plunger inserted into the rear of the cavity to adjust the cavity length while using a bellows to maintain a vacuum seal. In each measurement for a given pressure and pump power, the plunger was adjusted to tune the cavity resonance either to the direct or refilling transition, then the detuning of the CO<sub>2</sub> laser was adjusted with a piezoelectric transducer mounted on a pump cavity mirror to maximize the OPFIR laser output intensity. The 1.2mm wavelength laser radiation strongly diffracted from the 1mm diameter pinhole, so only a fraction was captured for detection through a series of copper pipes, copper mirrors, and Teflon lenses. Attenuators with unknown absorbance were used to limit the power reaching the detector to ensure it remained in the linear-response regime. A Virginia Diodes heterodyne receiver operating in the 220–330GHz band was used to measure the frequency and strength of the OPFIR laser emission. The width of the measured emission lines was less than 3MHz, determined not by the laser but by the detection bandwidth of the spectrum analyzer used to display the intermediate frequency.

**Theoretical model.** The infrared pump creates non-thermal rotational populations  $N_{\ell}^r$  and non-equilibrium populations of vibrational thermal pools  $N_p^v$ , and the theoretical goal is to predict these differences  $N$  from the equilibrium (unpumped) populations and their consequences for THz gain. Our semiclassical theoretical model consists of a set of rate equations giving  $\dot{N} = \partial N / \partial t$  for both  $N_{\ell}^r$  and  $N_p^v$ . Below, we begin by describing these rate equations (which include various

transitions, coupling to the pump, and an effective-temperature model for unknown vibrational levels), then explain how the  $\dot{N} = 0$  steady state is found, and finally how the output power is extracted self-consistently from the populations. Additional technical details are presented in SI.

The **rotational-level rate equations** for the symmetry type A being pumped are given by:

$$\begin{aligned}
\dot{N}_\ell^r(v, \mathbf{r}, t) = & \sum_{\ell'} -\gamma_{\ell\ell'} N_\ell^r(v, \mathbf{r}, t) + \gamma_{\ell'\ell} N_{\ell'}^r(v, \mathbf{r}, t) \\
& - (\gamma_{\ell p}^{\text{SPT}} + \gamma_{\ell p}^{\text{VS}}) N_\ell^r(v, \mathbf{r}, t) \\
& + R_{\text{pump}}(v, \mathbf{r}, t) \\
& + D\nabla^2 N_\ell^r(v, \mathbf{r}, t).
\end{aligned} \tag{1}$$

Here, population density  $N_\ell^r(v, \mathbf{r}, t)$  depends on molecular axial (the direction of the pump beam) velocity  $v$ , position  $\mathbf{r}$ , time  $t$ , and energy level  $\ell$ . Because the problem is axisymmetric, we simplify  $\mathbf{r}$  to only the radial position, eliminating the axial dependence by an averaging procedure described in SI. The first two terms in equation (1) describe dipole–dipole collisions, with  $\gamma_{\ell\ell'}$  denoting the transition rate from  $\ell \rightarrow \ell'$ , obtained from  $\gamma_{\ell\ell'} = n_{\text{tot}} \langle v_{\text{rel}} \rangle \sigma_{\ell\ell'}$ , where  $n_{\text{tot}}$  is the total population density,  $\langle v_{\text{rel}} \rangle$  is the thermally averaged relative velocity between molecules, and  $\sigma_{\ell\ell'}$  is the collision cross section measured to be  $320\text{\AA}^2$  between adjacent rotational levels.<sup>13</sup> The second line in equation (1) describes both a symmetry-preserving thermalization (SPT) process that moves non-thermal population in rotational level  $\ell$  to its associated vibrational thermal pool  $p$  with the same A or E symmetry type,<sup>11</sup> and a slower “V-swap” (VS) process that moves non-thermal rotational population to both A and E thermal pools with an equal probability.<sup>11</sup> Their cross sections are<sup>11,24</sup>  $\sigma^{\text{SPT}} = 137\text{\AA}^2$  and  $\sigma^{\text{VS}} = 21\text{\AA}^2$  at room temperature. The  $R_{\text{pump}}(v, \mathbf{r}, t)$  term involves

only levels L and U as the L→U transition rate induced by absorption of the pump laser:<sup>25</sup>

$$R_{\text{pump}}(v, \mathbf{r}, t) = \left[ N_{\text{L}}^{\text{tot}}(v, \mathbf{r}, t) - \frac{g_{\text{L}}}{g_{\text{U}}} N_{\text{U}}^{\text{tot}}(v, \mathbf{r}, t) \right] p_{\text{L} \rightarrow \text{U}} \cdot (\delta_{l, \text{U}} - \delta_{l, \text{L}}), \quad (2)$$

in which  $N_{\text{L}}^{\text{tot}}(v, \mathbf{r}, t)$  and  $N_{\text{U}}^{\text{tot}}(v, \mathbf{r}, t)$  are the sum of non-thermal and thermal population for levels L and U,  $g_{\text{L}}$  and  $g_{\text{U}}$  are the degeneracy of the two levels, and the transition probability is given by  $p_{\text{L} \rightarrow \text{U}} = \frac{8\pi^3}{3h^2c} \frac{|\langle J', K, V+1 | \mu | J, K, V \rangle|^2}{\pi R^2} [P_{\text{pump}}^{\text{forw}} f_b^{\text{forw}}(\nu, \nu_p) + P_{\text{pump}}^{\text{back}} f_b^{\text{back}}(\nu, \nu_p)]$ . Here,  $h$  is the Planck constant,  $c$  is the speed of light,  $R$  is the cell radius,  $|\langle J', K, V+1 | \mu | J, K, V \rangle|^2$  is the measured<sup>25</sup> dipole matrix element 0.027 Debye<sup>2</sup> for the pump transition,  $P_{\text{pump}}$  is the effective  $z$ -averaged pump power, and  $f_b(\nu, \nu_p)$  is the area-normalized lineshape function including pressure<sup>27</sup> and AC-Stark<sup>26</sup> broadening. Both forward (incident direction of pump beam, or  $+z$ ) and backward ( $-z$ ) directions are included due to the reflection of the back mirror. The formulae for  $P_{\text{pump}}$  and  $f_b(\nu, \nu_p)$  are further discussed in SI. The final term in equation (1) describes the spatial diffusion with a diffusion coefficient  $D = \frac{1}{3} \langle v_{\text{abs}} \rangle \ell_{\text{mf}}$  where  $\langle v_{\text{abs}} \rangle$  is the averaged absolute velocity and  $\ell_{\text{mf}}$  is the mean free path.<sup>1</sup> We discretized the diffusion equation with a second-order finite-difference approximation in cylindrical coordinates and with a spatial resolution  $10\mu\text{m}$  divided by pressure/Torr (since higher pressures have smaller  $\ell_{\text{mf}}$  that requires finer resolution), with boundary conditions described below. As mentioned above, we approximate the populations as uniform in  $z$  with an averaged pump power, which is accurate in the optimal regime where the pump intensity mostly “fills” the cavity, and allows us to solve only a 1d (radial) diffusion equation.

The **rate equations for the vibrational thermal pools** are dominated by collisions with the cell walls (which introduces a radial dependence) and inter-molecular collisions. For a symmetry

type  $S$  ( $= A$  or  $E$ ), they are:

$$\begin{aligned}
\dot{N}_{p,S}^v(\mathbf{r}, t) &= \sum_{p'} -\gamma_{pp'} N_{p,S}^v(\mathbf{r}, t) + \gamma_{p'p} N_{p',S}^v(\mathbf{r}, t) \\
&+ \dot{N}_{p,S}^{VS}(\mathbf{r}, t) \\
&+ \sum_{\ell} \left[ \delta_{S,A} \gamma_{\ell p}^{\text{SPT}} + \frac{\gamma_{\ell p}^{\text{VS}}}{2} \right] \int dv N_{\ell}^r(v, \mathbf{r}, t) \\
&+ D \nabla^2 N_{p,S}^v(\mathbf{r}, t).
\end{aligned} \tag{3}$$

(Unlike the rotational populations, we do not distinguish the vibrational levels  $N_{p,S}^v$  by axial velocity  $v$  since they do not directly absorb the pump.) The first two terms describe the intermolecular collision between vibrational level  $p \rightarrow p'$  and vice versa, for instance between  $V_3$  and  $V_6$  with a cross section  $\sigma_{V_6 \rightarrow V_3} = 3.21 \text{\AA}^2$ .<sup>11</sup> V-swap transitions also occur among vibrational states, described by the second line in equation (3). We consider the most common form of this collision between a  $V_3A$  and a  $V_0E$  molecule:  $V_3A + V_0E \leftrightarrow V_0A + V_3E$ .<sup>11</sup> The third line in equation (3) describes the SPT and VS processes between non-thermal rotational level  $\ell$  and the vibrational thermal pool  $p$ , as described in equation (1). The final term describes the diffusion process similar to above.

Although we use a standard Neumann boundary condition ( $\partial N / \partial r = 0$ ) for rotational levels to represent impermeable walls, we define a different vibrational-level boundary condition that expresses relaxation via molecule-wall collisions. Similar to previous work,<sup>1</sup> we assume that wall collisions “thermalize” the vibrational levels and redistribute molecules into all possible vibrational levels of the same symmetry  $S$  according to a Boltzmann distribution characterized by  $T_{\text{wall}}$ . Similar to “surface-reaction” boundary conditions used in chemistry,<sup>29</sup> we use a Robin boundary

condition:

$$-D \frac{\partial N_p}{\partial r} = \bar{v}(1 - q_p)N_p/2 - \bar{v}q_p \sum_{p' \neq p} N_{p'}/2, \quad \text{at } r = R \quad (4)$$

in which  $q_p$  is the population fraction of vibrational level  $p$  for a Boltzmann distribution at the wall temperature ( $T_{\text{wall}} = 300\text{K}$ ) and  $\bar{v}$  is the average velocity of the population flux perpendicular to the wall. A detailed derivation in SI shows  $\bar{v} = \langle v_{\text{abs}} \rangle / 2$ . There is an extra factor of 1/2 in equation (4) because only half of the population is moving towards the wall. (Note that the diffusion term  $D\nabla^2 N$  already captures motion of molecules to/from the wall, so the purely local boundary condition here is more correct than the nonlocal collision rate used in past work.<sup>1</sup>)

**Effective vibrational temperature (EVT) model.** As mentioned in the main text, more higher-lying vibrational levels must be included in the high-pressure regime. Although the energies of these vibrational levels are known,<sup>25</sup> the millions of relaxation processes among them are infeasible to measure. To solve this problem, we propose to model all accessible vibrational levels with one expandable thermal pool  $V_\Sigma$  which contains all 120 vibrational levels higher than  $V_3$  up to  $6959 \text{ cm}^{-1}$  ( $\sim 35k_B T_{\text{wall}}$ ), taking into account the specific energies and degeneracies of each. We argued in the main text that due to fast relaxation processes, all vibrational levels satisfy a Boltzmann distribution with a spatially varying temperature  $T_v$  that is determined self-consistently by the population density of  $V_0$  and  $V_3$  for both A and E symmetries:

$$T_v^S = -\frac{\Delta E}{k_B} \frac{1}{\log(N_{3,S}^{\text{tot}}/N_{0,S}^{\text{tot}})} \text{ where } \Delta E = E_3 - E_0 \text{ is the energy difference between } V_3 \text{ and } V_0.$$

In order to maintain the Boltzmann distribution for  $V_\Sigma$ , we introduce transition rates between  $V_3$  and  $V_\Sigma$  that satisfy  $\frac{\gamma_{3 \rightarrow \Sigma}}{\gamma_{\Sigma \rightarrow 3}} = \exp\left(\frac{E_3 - \sum_{k \in V_\Sigma} g_k E_k}{k_B T_v^S}\right)$ , and, by taking the limit  $\gamma_{3 \rightarrow \Sigma}, \gamma_{\Sigma \rightarrow 3} \rightarrow \infty$ , we obtain  $N_{\Sigma,S}(r) = \exp\left(\frac{E_3 - \sum_{k \in V_\Sigma} g_k E_k}{k_B T_v^S}\right) N_{3,S}(r)$  for the steady state.

**Numerical Solver.** The nonlinear rate equations are solved directly by setting  $\dot{N}_\ell = 0$  and  $\dot{N}_p = 0$  in Eqs. (1, 3). The nonlinearity of the rate equations arises from the pump term, V-swap collision between vibrational levels, and the EVT model. Our nonlinear solver is based on fixed-point iteration with Anderson acceleration,<sup>30</sup> which is convenient because it does not require us to compute an explicit Jacobian matrix and makes it easy to add new physics and modify the  $\dot{N}$  equations. Typically, we have  $\approx 5 \times 10^5$  degrees of freedom (20 energy levels  $\times$  50 radial points  $\times$  500 velocity subclasses). The solution typically converges within 20 Anderson iterations.

**Output Power.** The lasing output power can be obtained by matching terahertz absorption  $\alpha_{\text{THz}}$  with the gain:

$$\alpha_{\text{THz}} = \frac{\int_V \gamma(\nu, r) |\mathbf{E}(r)|^2 dr}{\int_V |\mathbf{E}(r)|^2 dr}, \quad (5)$$

where  $\alpha_{\text{THz}}$  is the cavity loss, including ohmic loss and leakage through the front window. Ohmic loss can be computed analytically for the modes of a hollow metal waveguide,<sup>28</sup> and the lowest-loss mode is TE<sub>01</sub>, but which mode lases will depend on the degree of radial and frequency overlap between the gain and mode profiles. Since the lasing mode pattern could not be observed directly in the experiment, we used the  $\alpha_{\text{THz}}$  from TE<sub>01</sub> for all curves except for refilling lasing with input pump power 6.6W. The 6.6W experimental refilling curve exhibited a jump in amplitude compared to adjacent powers, and we believe that this is due to a different mode being excited by a mistuned cavity; using  $\alpha_{\text{THz}}$  and the mode profile from TE<sub>22</sub> best fit the data, indicating that the model could enable us to determine which mode is lasing without directly measuring the mode profile! Since the gain is a small perturbation ( $< 0.01\%$ ) to the permittivity, the electric field  $\mathbf{E}(r)$  is accurately approximated by the passive waveguide mode. For convenience, we normalize the mode so that

$\int_V |\mathbf{E}(r)|^2 d\mathbf{r} = \int_V d\mathbf{r} = V$ . The gain coefficient  $\gamma(\nu, r)$  at lasing frequency  $\nu$  and radial position  $r$  is then expressed as<sup>27</sup>

$$\gamma(\nu, r) = \int dv \frac{\Delta N(v, r) \sigma(\nu)}{1 + \Phi |\mathbf{E}(r)|^2 / \Phi_s} \quad (6)$$

which includes contributions from all velocity subclasses.  $\Delta N(v, r)$  is the population inversion obtained by solving the rate equations, and  $\sigma(\nu)$  is the transition cross section:  $\sigma(\nu) = \lambda^2 g(\nu) / 8\pi t_{\text{sp}}$ , where  $\lambda$  is the lasing wavelength, the spontaneous emission lifetime  $t_{\text{sp}}$  is taken to be 10sec,<sup>1</sup> much longer than all other transition lifetimes in the considered system (see SI for a detailed discussion:  $t_{\text{sp}}$  has little effect on the peak efficiency), and  $g(\nu) = \frac{1}{\pi} \frac{\Delta\nu_P}{(\nu - \nu_v)^2 + \Delta\nu_P^2}$  is the pressure-broadening lineshape function ( $\nu_v$  is the center frequency for molecules with velocity  $v$ ).  $\Phi_s$  is the saturated photon flux density:  $\Phi_s = 1/\tau_s \sigma(\nu)$ , where  $\tau_s$  is the characteristic time for the lasing transition. Unlike a three- or four-level system in which  $\tau_s$  can be analytically derived,<sup>27</sup> our system requires  $\tau_s$  to be obtained numerically from the model. A stimulated emission rate  $w_{\text{st}}$  is introduced between the two lasing levels, and  $\tau_s$  is obtained by fitting the population inversion as a function of  $w_{\text{st}}$ :  $\Delta N(w_{\text{st}}) = \Delta N(w_{\text{st}} = 0) / (1 + \tau_s w_{\text{st}})$ . Once all  $\Delta N$ ,  $\sigma$ ,  $\Phi_s$  and cavity mode  $\mathbf{E}(r)$  are known, we numerically solve equation (5) for the intensity of photon flux density  $\Phi$  by a Newton method. Finally, the THz output power is obtained from  $I_o = h\nu AT\Phi/2$ , where  $A$  is the cross section of the cavity and  $T$  is the transmission coefficient, computed by integrating the waveguide mode's Poynting flux over the output pinhole (since the pinhole is comparable to the wavelength; see SI). The formula is divided by 2 to count only photons flowing in the outward direction.<sup>27</sup>

1. Chua, S.-L. *et al.* Spatio-temporal theory of lasing action in optically-pumped rotationally excited molecular gases. *Optics Express* **19**, 7513–7529 (2011).
2. Dangoisse, D., Glorieux, P. & Wascat, J. Diffusion and vibrational bottleneck in optically pumped submillimetre lasers. *International Journal of Infrared and Millimeter Waves* **2**, 215–229 (1981).
3. Henningsen, J. & Jensen, H. The optically pumped far-infrared laser: Rate equations and diagnostic experiments. *IEEE Journal of Quantum Electronics* **11**, 248–252 (1975).
4. Hodges, D. T., Tucker, J. R. & Hartwick, T. S. Basic physical mechanisms determining performance of the CH<sub>3</sub>F laser. *Infrared Physics* **16**, 175–182 (1976).
5. Chang, T. Y. & Bridges, T. J. Laser action at 452, 496, and 541  $\mu\text{m}$  in optically pumped CH<sub>3</sub>F. *Optics Communications* **1**, 423–426 (1970).
6. Matteson, W. & De Lucia, F. Millimeter wave spectroscopic studies of collision-induced energy transfer processes in the <sup>13</sup>CH<sub>3</sub>F laser. *IEEE Journal of Quantum Electronics* **19**, 1284–1293 (1983).
7. Tobin, M. S. A review of optically pumped NMMW lasers. *Proc. IEEE* **73**, 61–85 (1985).
8. Everitt, H. O., Skatrud, D. D. & De Lucia, F. C. Dynamics and tunability of a small optically pumped cw far-infrared laser. *Applied Physics Letters* **49**, 995–997 (1986).

9. McCormick, R. I., De Lucia, F. C. & Skatrud, D. D. A time-resolved study of rotational and vibrational excitation and relaxation in the  $^{13}\text{CH}_3\text{F}$  optically pumped far-infrared laser. *IEEE Journal of Quantum Electronics* **23**, 2060–2068 (1987).
10. McCormick, R. I., Everitt, H. O., De Lucia, F. C. & Skatrud, D. D. Collisional energy transfer in optically pumped far-infrared lasers. *IEEE Journal of Quantum Electronics* **23**, 2069–2077 (1987).
11. Everitt, H. O. & De Lucia, F. C. A time-resolved study of rotational energy transfer into A and E symmetry species of  $^{13}\text{CH}_3\text{F}$ . *Journal of Chemical Physics* **90**, 3520–3527 (1989).
12. Douglas, N. G. *Millimetre and Submillimetre Wavelength Lasers: A Handbook of CW Measurements*, vol. 61 (Springer, 1989).
13. Everitt, H. O. & De Lucia, F. C. Rotational energy transfer in  $\text{CH}_3\text{F}$ : The  $\Delta J = n$ ,  $\Delta K = 0$  processes. *The Journal of Chemical Physics* **92**, 6480–6491 (1990).
14. Tonouchi, M. Cutting-edge terahertz technology. *Nature photonics* **1**, 97–105 (2007).
15. Ferguson, B. & Zhang, X.-C. Materials for terahertz science and technology. *Nature materials* **1**, 26–33 (2002).
16. Siegel, P. H. Terahertz technology. *IEEE Transactions on Microwave Theory and Techniques* **50**, 910–928 (2002).
17. Edinburgh Instruments Ltd. <https://www.edinst.com/products/>.

18. Button, K. J. *Reviews of Infrared and Millimeter Waves: Volume 2 Optically Pumped Far-Infrared Lasers*, vol. 2 (Springer Science & Business Media, 2013).
19. Johnson, A. & Auston, D. Microwave switching by picosecond photoconductivity. *IEEE Journal of Quantum Electronics* **11**, 283–287 (1975).
20. Van Exter, M. & Grischkowsky, D. Optical and electronic properties of doped silicon from 0.1 to 2 THz. *Applied Physics Letters* **56**, 1694–1696 (1990).
21. Andronov, A. A. *et al.* The gyrotron: High-power source of millimetre and submillimetre waves. *Infrared Physics* **18**, 385–393 (1978).
22. Faist, J. *et al.* Quantum cascade laser. *Science* **264**, 553–555 (1994).
23. Williams, B. S. Terahertz quantum-cascade lasers. *Nature Photonics* **1**, 517–525 (2007).
24. Everitt, H. O. & De Lucia, F. C. Rotational energy transfer in small polyatomic molecules. *Advances in Atomic, Molecular, and Optical Physics* **35**, 331 (1995).
25. Everitt III, H. O. *Collisional Energy Transfer in Methyl Halides*. Ph.D. thesis, Duke University (1990).
26. Panock, R. & Temkin, R. Interaction of two laser fields with a three-level molecular system. *IEEE Journal of Quantum Electronics* **13**, 425–434 (1977).
27. Saleh, B. E. A. & Teich, M. C. *Fundamentals of Photonics*, vol. 22 (Wiley, 1991).
28. Bansal, R. *Fundamentals of Engineering Electromagnetics* (CRC press, 2006).

29. Fogler, H. S. *Elements of Chemical Reaction Engineering* (Prentice-Hall International London, 1999).
30. Walker, H. F. & Ni, P. Anderson acceleration for fixed-point iterations. *SIAM Journal on Numerical Analysis* **49**, 1715–1735 (2011).

**Acknowledgements** We thank Prof. Frank De Lucia, Prof. Richard Temkin, Dr. Ling Lu, Dr. Bo Zhen, and Dr. Adi Pick for helpful discussions. This work was supported in part by the U. S. Army Research Laboratory and the U. S. Army Research Office through the Institute for Soldier Nanotechnologies under contract number W911NF-13-D-0001, and by the MRSEC Program of the National Science Foundation under award number DMR-1419807. J.B.-A. also acknowledges financial support by Spanish MINECO under grant MAT2015-66128-R (MINECO/FEDER).

**Author contributions** H.O.E., M.S., J.D.J., and S.G.J. conceived the idea for this study and coordinated the work, with H.O.E. and S.G.J. providing close supervision. F.W., J.L., S.L.C., and J.B.-A. developed the theoretical model and the numerical solver. D.J.P. and S.G.H. performed the experimental measurements. F.W., S.G.J., and H.O.E. wrote the paper with input from all authors.

**Author Information** The authors declare that they have no competing financial interests. Correspondence and requests for materials should be addressed to F.W. (email: wangfan@mit.edu), S.G.J. (email: stevenj@math.mit.edu), and H.O.E. (email: everitt@phy.duke.edu).

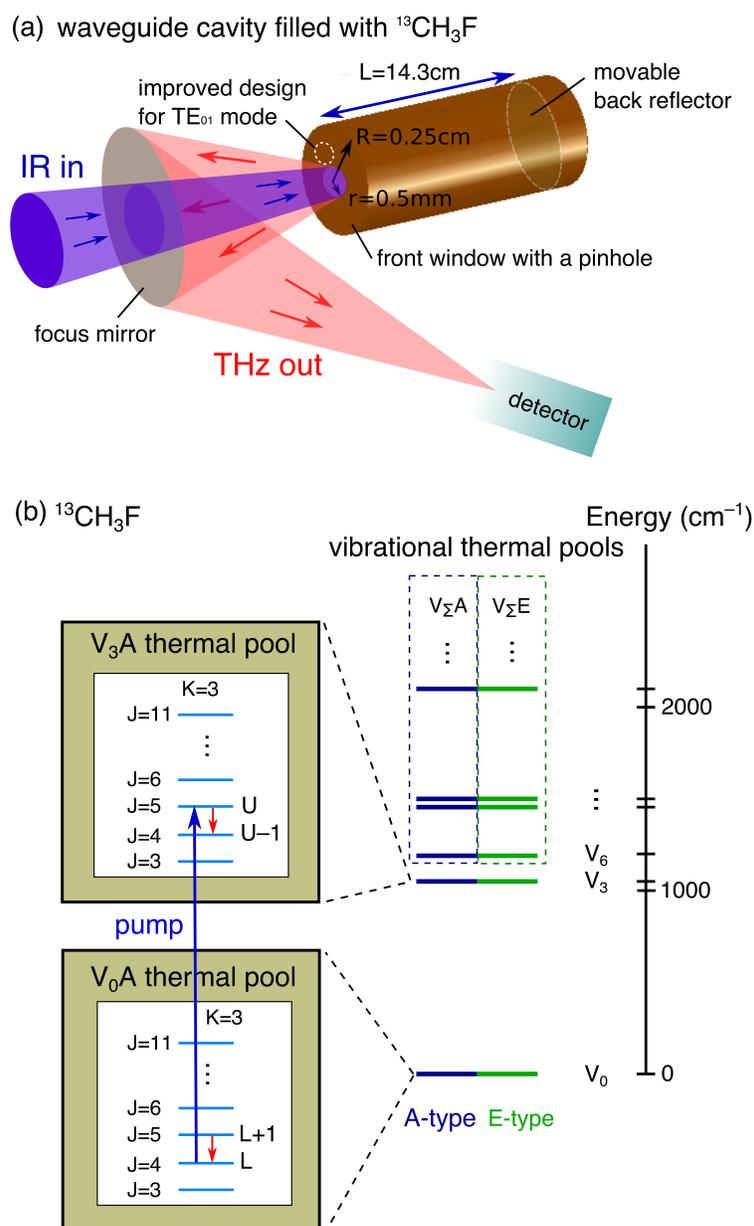


Figure 1: Schematics of (a) the laser cavity and (b) molecular energy levels of  $^{13}\text{CH}_3\text{F}$ . The cavity is a copper tube with a movable back wall used to tune the cavity frequency to match the laser gain, pumped with an IR laser through a pinhole in the front window. The energy levels consist of vibrational levels  $V_\ell$ , with A and E symmetries,<sup>11</sup> composed of rotational levels with quantum numbers J and K. The IR pump excites a transition between levels L and U as shown, and lasing occurs for both the “direct” inversion between U and U-1 and the “refilling” inversion between L+1 and L.

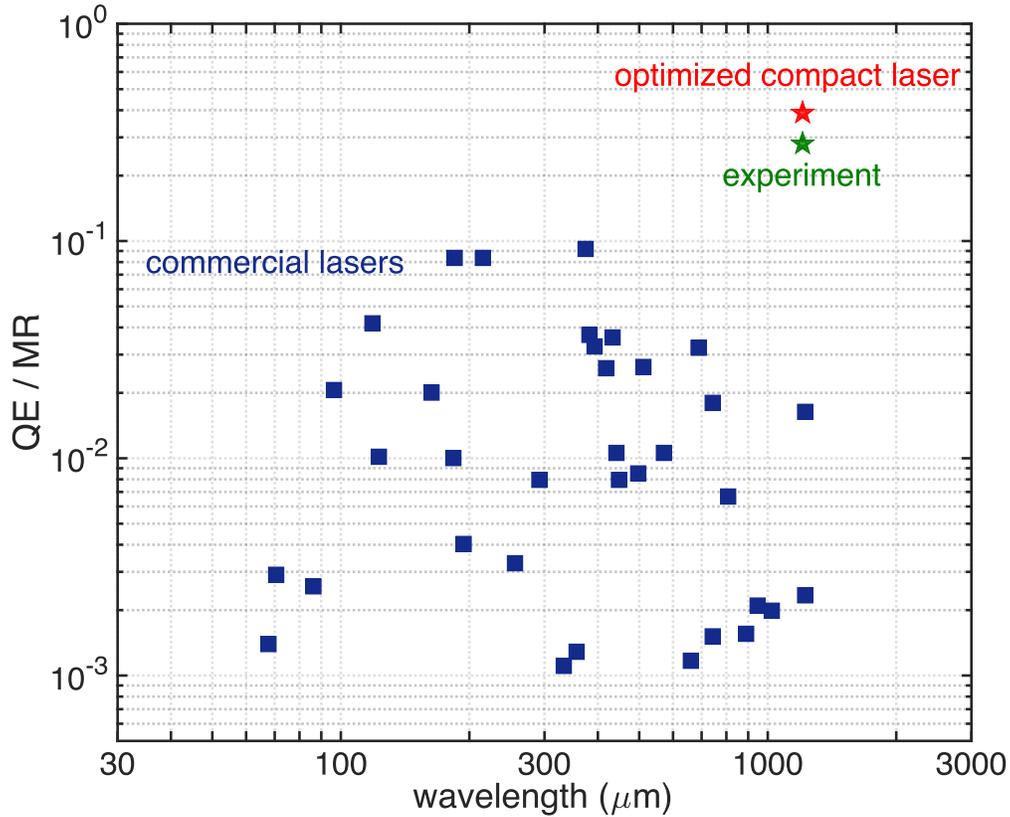


Figure 2: Total quantum efficiency<sup>27</sup> ( $QE = \text{THz power out} / \text{IR power in}$ ) of commercial OPFIR lasers<sup>17</sup> and our compact OPFIR laser, normalized by the Manley–Rowe (MR) limit<sup>27</sup> on QE. Our experimentally demonstrated laser achieves a QE that is 29% of the MR limit (29% of 0.8%) which improves to 39% after cavity optimization. Both are  $10\times$  better than the best commercial laser at the same frequency (0.25THz, or 1.2mm wavelength), while being  $1000\times$  smaller—in fact, we show theoretically that this efficiency boost is mainly due to the fact that our cavity is so much smaller.

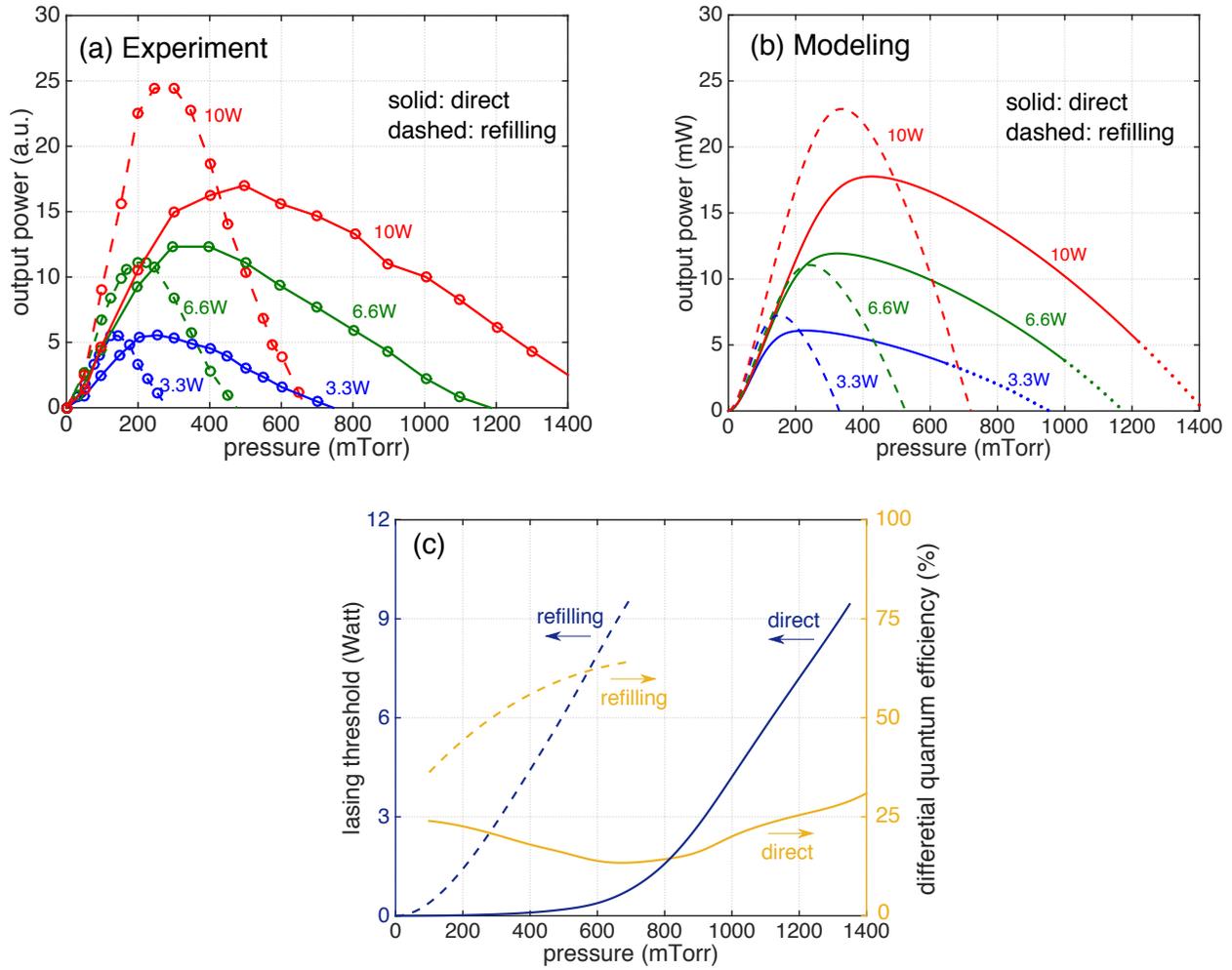


Figure 3: (a) Experimental and (b) theoretical THz laser output power as a function of gas pressure, for various input pump powers (3.3W, 6.6W, and 10W), and for both the direct (solid lines) and refilling (dashed lines) transitions. Quantitative agreement is obtained even though the theoretical model has no adjustable parameters; in contrast to previous work,<sup>1-4</sup> our theory captures the fact that higher powers allow operation at higher pressures, in part because of a novel technique to model a large number of high-energy vibrational levels. The dotted parts of the direct lasing curves represent an inefficient regime in which the pump intensity is nearly zero in the back of the cavity. In this case our model's approximation of an effective  $z$ -uniform pump strength is inaccurate, but it is not desirable to operate in this regime. Our model also predicts (c) the input pump power at lasing threshold (blue) and the differential quantum efficiency (DQE, yellow)<sup>27</sup> normalized by the MR limit at threshold versus pressure for both inversions. The refilling inversion is both higher power and higher efficiency, with a modeled DQE that can reach 64%.

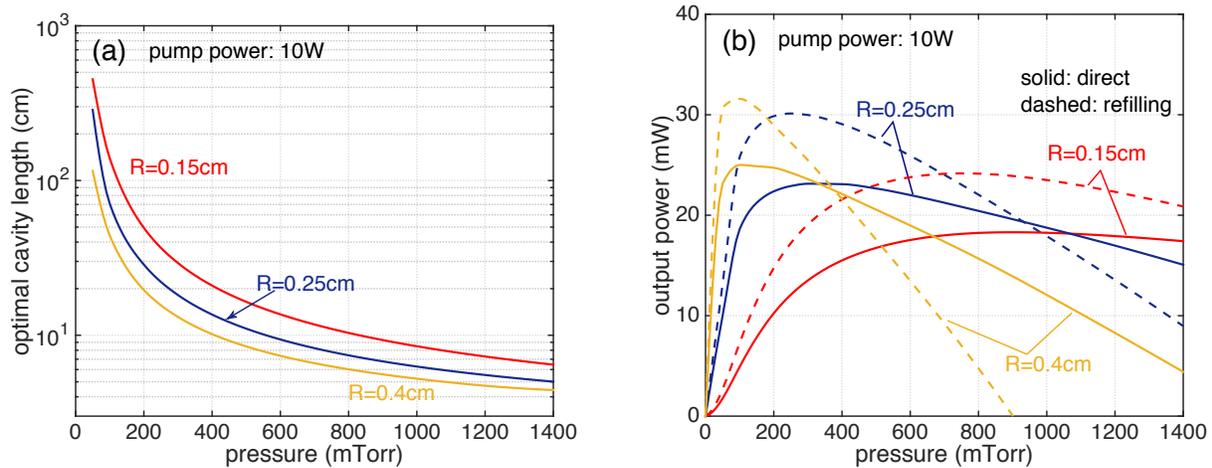


Figure 4: Theoretical predictions of (a) the roughly optimal cavity length for various radii and (b) the optimized laser power, both as a function of pressure for an input pump power of 10W. The cavity length is optimized to equal half the decay length of the pump beam. The output power (b) is shown for the optimal cavity length at each pressure and radius so that each data point is a separate cavity design. The peak efficiency is increased from 29% of Manley–Rowe in the experimental cavity to 39%, but more importantly this optimization allows efficient operation at a wider range of high pressures and hence in smaller cavities.

# Supplementary Information

## 1 Broadening linewidth

As explained in Methods,  $f_b(\nu, \nu_p)$  is the area-normalized lineshape function (including pressure<sup>1</sup> and AC-Stark<sup>2</sup> broadening),  $\nu_p$  is the pump frequency, and  $\nu$  is the transition frequency of molecules with axial velocity  $v$ . By the Doppler effect,  $\nu^\pm = (1 \pm v/c)\nu_0$  for the forward (+) and backward (−) directions, where  $\nu_0 = 31.04\text{THz}$  is the intrinsic transition frequency between L and U for  $^{13}\text{CH}_3\text{F}$ . The lineshape function  $f_b^\pm(\nu^\pm, \nu_p)$  is a convolution of pressure and AC-Stark broadening effects:

$$f_b^\pm(\nu^\pm, \nu_p) = f_p(\nu^\pm, \nu_p) * f_S^\pm(\nu^\pm, \nu_p)$$

where

$$f_p(\nu^\pm, \nu_p) = \frac{1}{\pi} \frac{\Delta\nu_P}{(\nu^\pm - \nu_p)^2 + \Delta\nu_P^2}$$

is the pressure broadening (with linewidth  $\Delta\nu_P = \frac{1}{2\pi\Delta\tau} \approx 15\text{MHz/Torr}$  with  $\Delta\tau$  being the mean collision time<sup>3</sup>), and the AC-Stark broadening  $f_S^\pm(\nu^\pm, \nu_p)$  is also approximately Lorentzian with linewidth<sup>2</sup>

$$\Delta\nu_S^\pm = \frac{|\langle J', K, V + 1 | \mu | J, K, V \rangle|}{2\pi\hbar} \sqrt{\frac{2P_{\text{pump}}^\pm}{\pi\epsilon_0 c R}} = 0.38 \sqrt{\frac{P_{\text{pump}}^\pm/\text{Watt}}{R/\text{cm}}}.$$

(Note that the average pump power  $P_{\text{pump}}^\pm$  is different for the  $+z$  direction and the reflected power in the  $-z$  direction, as described below.) The convolution of two Lorentzians is also a Lorentzian

with the sum of the linewidths, so we obtain

$$f_b^\pm(\nu^\pm, \nu_p) = \frac{1}{\pi} \frac{\Delta\nu_b}{(\nu^\pm - \nu_p)^2 + (\Delta\nu^\pm)^2}$$

with

$$\Delta\nu^\pm = \Delta\nu_p + \Delta\nu_S^\pm.$$

## 2 Pump absorption coefficient $\alpha$ and the average pump power

The pump absorption coefficient as a function of radial position  $r$  is<sup>3</sup>

$$\alpha^\pm(r) = \frac{8\pi^3\nu_p}{3hc} |\langle J', K, V + 1 | \mu | J, K, V \rangle|^2 \int g^\pm(v) \left[ N_L^{\text{tot}}(v, r) - \frac{g_L}{g_U} N_U^{\text{tot}}(v, r) \right] dv$$

in which  $h$  is the Planck constant,  $c$  is the speed of light,  $R$  is the cavity radius,  $\nu_p$  is the pump frequency, and  $|\langle J', K, V + 1 | \mu | J, K, V \rangle|^2$  is the measured<sup>3</sup> dipole matrix element 0.027 Debye<sup>2</sup> for the pump transition. In the integral over velocity,  $g^\pm(v) = f_b^\pm(\nu^\pm(v), \nu_p)$  denotes the pump absorption spectrum for molecules with axial speed  $v$ , so that  $g^\pm(v)$  is larger if the pump frequency is closer to the Doppler-shifted L $\rightarrow$ U transition frequency at that  $v$ . Note that  $\alpha^\pm(r)$  is different in forward (+) and backward (−) directions since they have different Doppler shifts and powers. The spatially averaged pump absorption coefficient is obtained by:

$$\alpha^\pm = \frac{\int_0^R \alpha^\pm(r) r dr}{\int_0^R r dr}.$$

Note that  $\alpha_p$  in the main text for the definition of optimal cavity length ( $L_o = 1/2\alpha_p$ ) is the same with  $\alpha^+$ . Figure S1 shows IR absorption spectrum  $\alpha^+$  in the forward direction at 100 mTorr and 400 mTorr with various input pump powers.  $\nu_0 = 31.0427$  THz is the frequency of the L–U transition, while the CO<sub>2</sub> pump laser is incident at a frequency 25 MHz  $>$   $\nu_0$ . The figure shows a dip at the pump frequency with nonzero pump power because of reduced population in the L state. The width of the dip increases with pressure and pump power, as described in the previous section, becoming so broad at 400 mTorr that its effect is hard to see. There is an extra dip at  $-25$  MHz due to the Doppler-shifted pump reflection at the back mirror of the cavity.

Since the pump photon can travel back and forth inside the cavity, reflecting off of the front

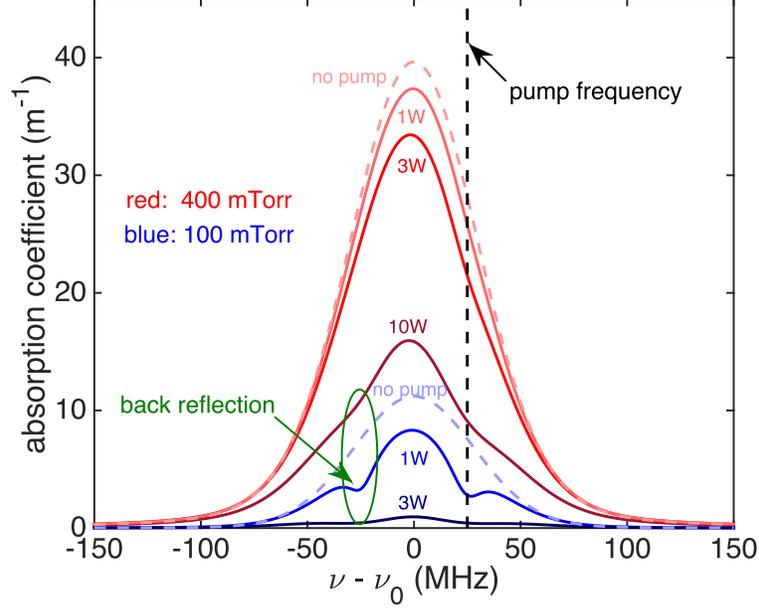


Figure 1: IR absorption spectrum at 100 mTorr and 400 mTorr with various input pump powers.  $\nu_0 = 31.0427$  THz is the frequency of the L–U transition, while the CO<sub>2</sub> pump laser is incident at a frequency 25 MHz  $>$   $\nu_0$ .

and back gold-coated copper mirrors, we sum all the trips to obtain an effective  $z$ -averaged pump power in each direction. The loss for one round trip is  $\beta = e^{-\alpha^+L - \alpha^-L} R_1 R_2$ , where  $R_2 = 0.95$  is the reflection on the back mirror at IR frequency 31.04 THz incident on a gold-coated copper plate,  $R_1 = 0.96 \times 0.95$  is the reflection of the front mirror considering an extra 4% leakage from the pinhole. Then the average pump power is computed by a  $z$ -averaging procedure:

$$P_{\text{pump}}^+ = \frac{P_0}{1 - \beta} \frac{\int_0^L \exp(-\alpha^+ z) dz}{L} = \frac{P_0}{1 - \beta} \frac{1 - \exp(-\alpha^+ L)}{\alpha^+ L}$$

$$P_{\text{pump}}^- = \frac{P_0 R_2 \exp(-\alpha^+ L)}{1 - \beta} \cdot \frac{1 - \exp(-\alpha^- L)}{\alpha^- L}$$

where  $P_0$  is the input pump power. The average pump power decreases as pressure and IR absorption coefficient increase for a given input pump power.

### 3 Artificial bottlenecking at high pressures

As mentioned in the main text, the reason why previous models produced artificial bottlenecking at high pressures was conjectured to be their inclusion of only a few vibrational levels. Here, we demonstrate this artificial bottlenecking phenomena by including all the rotational levels as shown in Fig. 1 of the main text, but only the 6 vibrational levels that are most directly connected to the lasing energy levels. In this 6-level model, we used measured transition rates between the vibrational levels, and all of the collisional and diffusion processes mentioned in the main text were included. The *only* difference is that it has no effective temperature to model the “missing” higher-lying vibrational levels, and the population of all 6 vibrational levels is solved exactly. The output THz powers with different input pump powers (3.3W, 6.6W, and 10W) are shown in Fig. S2. A clear pressure bottlenecking is seen around 700 mTorr, above which no lasing is achieved. This confirms the long-held understanding that including only few vibrational levels in the model induces artificial high-pressure quenching of the inversion through vibrational bottlenecking.

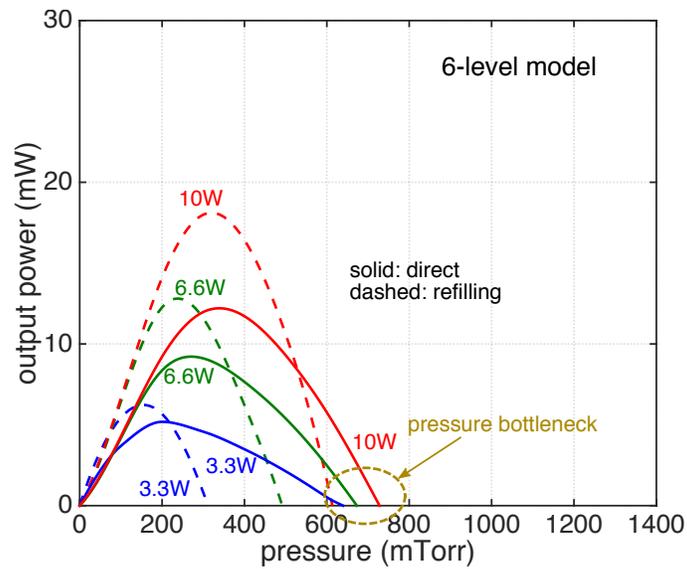


Figure 2: THz output power as a function of gas pressure, for various input pump powers (3.3W, 6.6W, and 10W), and for both the direct (solid lines) and refilling (dashed lines) transitions by a 6-level model. It is clearly observed that no gain is obtained above 700 mTorr by the 6-level model, known as high-pressure artificial bottlenecking.

#### 4 Radial distribution of population inversion and effective temperature $T_v$

The radial distribution of the population inversion  $\Delta N_U$  for the direct transition at 100 mTorr, 200 mTorr, 400 mTorr, and 1 Torr with 10W input pump power is plotted in Fig. S3. Since the primary pathway for vibrational relaxation is molecule–wall collisions, the population inversion is larger at radial positions closer to the cavity wall, as shown in the figure. This effect is crucial for cavity design.

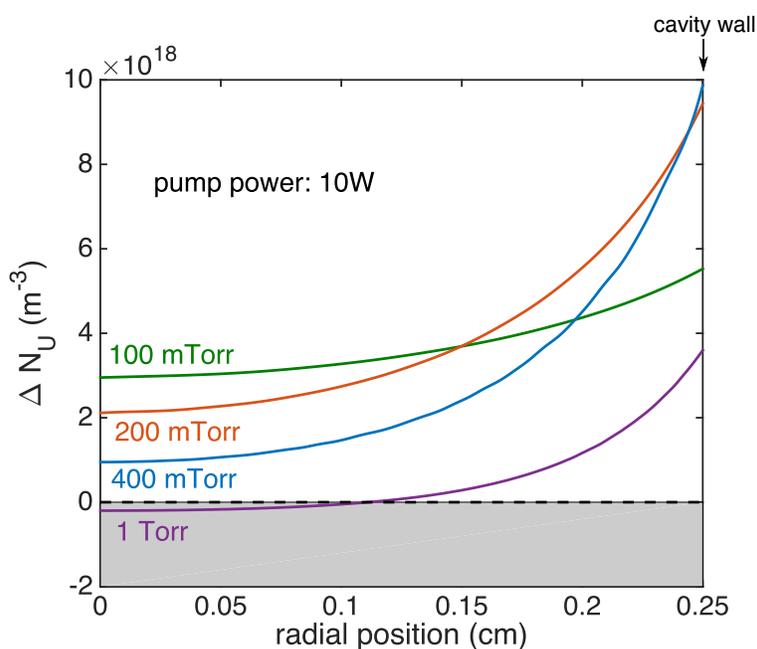


Figure 3: Radial distribution of population inversion for direct transition at 100 mTorr, 200 mTorr, 400 mTorr, and 1 Torr, with 10W input pump power.

The effective vibrational temperature  $T_v$  provides a deeper understanding of the system dynamics. Figure S4 shows a contour plot of  $T_v$  for the A-type vibrational thermal pools as a function of radial position and molecular pressure, for an input pump power of 10W and a cavity length

$L = 14.3\text{cm}$ . The effective vibrational temperature is higher in the cell center, where the pump is more saturated and the gain is smaller. One would normally expect  $T_v$  to increase with pressure due to faster intermolecular collisions, but in this figure  $T_v$  actually *decreases* above 400 mTorr because the  $z$ -averaging of the pump power lowers the pump rate and brings the effective temperature down. If we plot the  $T_v$  with optimal cavity length (which decreases with pressure so that the whole cavity remains pumped) as discussed in the main text, the result (shown in Fig. S5) exhibits effective temperatures that increase with pressure as expected.

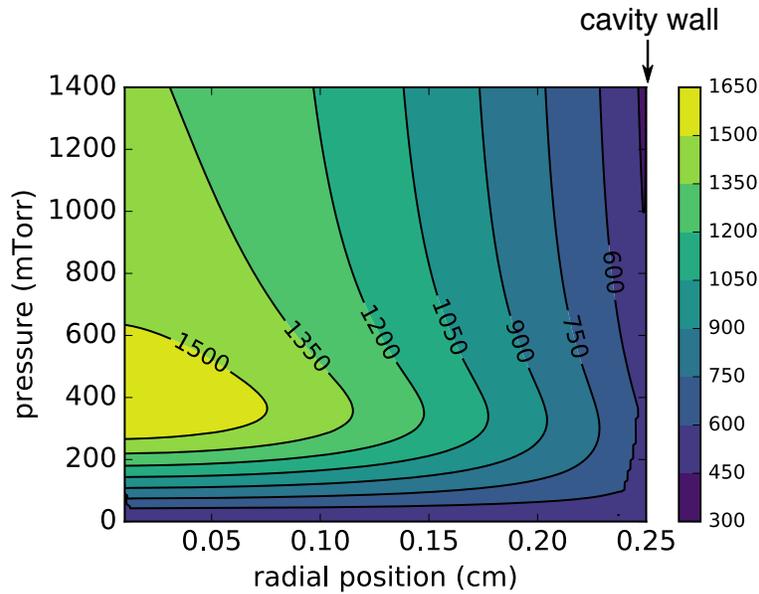


Figure 4: The effective temperature  $T_v$  (in Kelvin) of type-A vibrational thermal pools as a function of radial position and molecular pressure, for an input pump power of 10W. The cavity is  $5\text{mm} \times 14.3\text{cm}$  as in experiments.

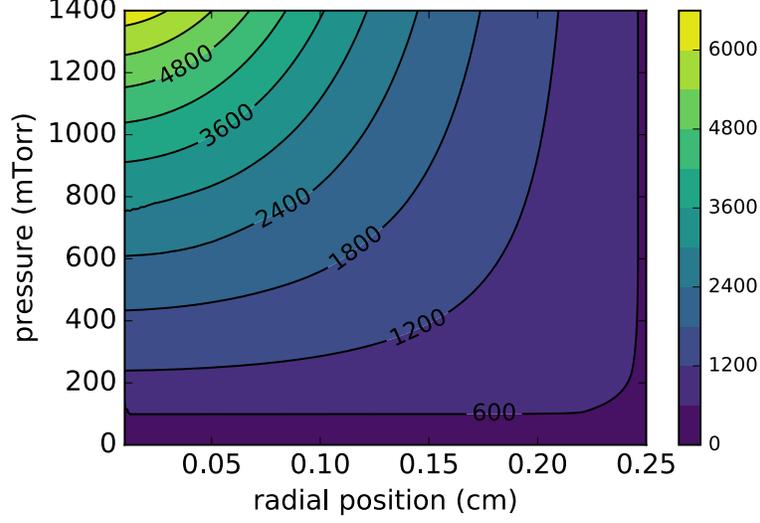


Figure 5: The effective temperature  $T_v$  (in Kelvin) of type-A vibrational thermal pools as a function of radial position and molecular pressure, for an input pump power of 10W. The cavity length is optimized to match the IR absorption:  $L_o = 1/2\alpha_p$ , and the cavity diameter is 5mm.

## 5 THz gain spectrum

The small-signal gain spectrum is computed as

$$\gamma_0(\nu) = \int dv \langle \Delta N(v) \rangle \sigma(\nu)$$

where  $\langle \Delta N(v) \rangle$  is the spatially averaged population inversion of velocity subclass  $v$ , and  $\sigma(\nu)$  is the transition cross section (see Methods). Typical gain spectra are plotted in Fig. S6 at 100 mTorr and 400 mTorr with the input pump power equal to 3W and 10W. The width of THz gain spectrum is proportional to the gas pressure due to pressure broadening.

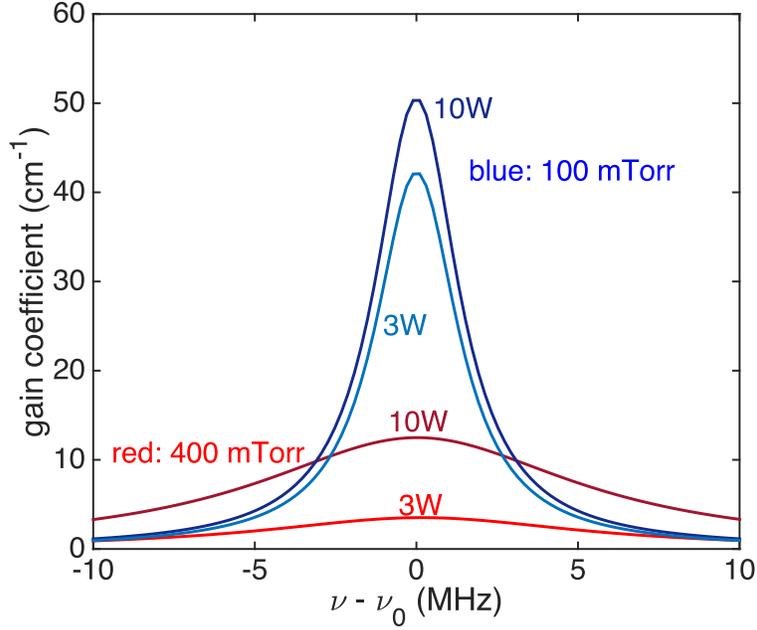


Figure 6: THz gain spectrum at 100 mTorr and 400 mTorr with the input pump power equal to 3W and 10W. The width of spectrum is proportional to the gas pressure due to pressure broadening.

## 6 Spontaneous-emission lifetime

To compute the gain coefficient and the output power, the spontaneous emission lifetime  $t_{\text{sp}}$  is taken to be 10 seconds in Methods. In fact, the precise value of  $t_{\text{sp}}$  has little effect on the peak output power far above threshold. In this section, we first demonstrate the effect of  $t_{\text{sp}}$  by explicit numerical calculations (re-running our model with different values of  $t_{\text{sp}}$ , shown in Figure S7), and then we present a simple analytical explanation for the lack of influence of  $t_{\text{sp}}$ .

Figure S7 shows the numerical output power of both direct (solid lines) and refilling (dashed lines) lasing as a function of pressure with the input pump power equal to 10W, and with three different spontaneous emission lifetimes  $t_{\text{sp}}$ : 1sec (blue), 10sec (red), and 100sec (green). The

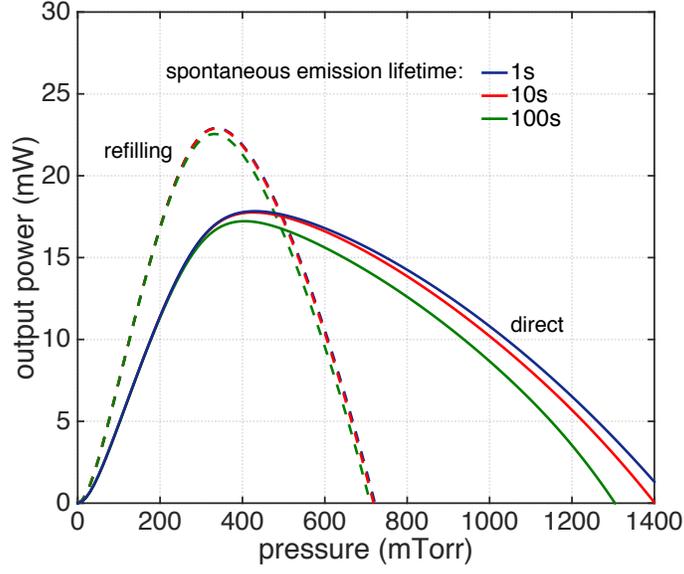


Figure 7: Numerical output power of both direct (solid lines) and refilling (dashed lines) lasing versus pressure with the input pump power equal to 10W, and with three different spontaneous emission lifetimes  $t_{sp}$ : 1sec, 10sec, and 100sec.  $t_{sp}$  has little effect on the peak output power far above threshold for both lasing transitions.

output power curves are very close, especially in the high-power regions far above the threshold, for both direct and refilling inversions. The biggest effect of  $t_{sp}$  is on the high-pressure cutoff for the direct transition, but in this (uninteresting, low-efficiency) high-pressure regime our model's  $z$ -averaging assumption is inaccurate as discussed in the main text.

It is straightforward to show analytically that the effect of  $t_{sp}$  approximately cancels in the high-power regime. In that regime, far above threshold,  $\Phi \gg \Phi_s$ , and Eq. (5) from the main text

becomes

$$\alpha_{\text{THz}} \approx \frac{1}{V} \int \int \frac{\Delta N(v) \sigma(\nu)}{\Phi \tau_s \sigma(\nu)} dv d\mathbf{r} = \frac{\langle \Delta N \rangle}{\Phi \tau_s}$$

where  $\langle \Delta N \rangle$  is the spatially averaged population inversion and  $\sigma$  is the transition cross section (which depends on  $t_{\text{sp}}$ ). This equation yields an approximate flux density  $\Phi \approx \langle \Delta N \rangle / \tau_s \alpha_{\text{THz}}$ , in which  $\sigma$  has cancelled and therefore the dependence on  $t_{\text{sp}}$  has vanished.

## 7 Average velocity of the population flux into the wall

The boundary condition for the vibrational levels, discussed in Methods, is written in terms of  $\bar{v}$ , the average molecule speed in the direction perpendicular to the wall (say,  $+x$ ). In this section, we briefly review the computation of  $\bar{v}$  by standard methods. In particular, the molecule velocities follow a Maxwell–Boltzmann distribution

$$f_{\text{MB}}(v_x) = \sqrt{\frac{m}{2\pi k_B T}} e^{-\frac{mv_x^2}{2k_B T}}$$

By averaging the velocities in the  $+x$  direction, one obtains

$$\begin{aligned}\bar{v} &= \frac{\int_0^\infty v_x f_{\text{MB}}(v_x) dv_x}{\int_0^\infty f_{\text{MB}}(v_x) dv_x} = 2 \int_0^\infty v_x f_{\text{MB}}(v_x) dv_x \\ &= \sqrt{\frac{m}{2\pi k_B T}} \int_0^\infty e^{-\frac{mv_x^2}{2k_B T}} d(v_x^2) = \sqrt{\frac{2k_B T}{\pi m}}.\end{aligned}$$

On the other hand, the average absolute velocity is

$$\langle v_{\text{abs}} \rangle = \left( \frac{m}{2\pi k_B T} \right)^{3/2} \int |\mathbf{v}| \exp\left(-\frac{mv^2}{2k_B T}\right) d\mathbf{v} = \sqrt{\frac{8k_B T}{\pi m}}.$$

Therefore  $\bar{v} = \langle v_{\text{abs}} \rangle / 2$ .

## 8 Transmission through the output pinhole

THz radiation transmitted through the output pinhole affects the calculation of total loss, lasing threshold, and the output power. Instead of using the ratio of pinhole area and the cavity cross section,<sup>1</sup> which is only correct for a uniform intensity (e.g. a plane wave), the transmission coefficient  $T$  is computed by integrating the waveguide mode's Poynting flux over the output pinhole. (This is also an approximation, which is valid because the 1mm diameter pinhole is not small compared to 1.2mm wavelength. The radiation through a very subwavelength hole would require an entirely different approach.<sup>4</sup>) For a circular waveguide's  $\text{TE}_{nm}$  mode,<sup>5</sup>

$$P_{\text{rad}} = \int_S \frac{1}{2} \text{Re}(\mathbf{E} \times \mathbf{H}^*)_z d\mathbf{r} = \frac{Z_{nm}^{\text{TE}}}{2} \int_S (|H_r|^2 + |H_\phi|^2) d\mathbf{r}$$

in which  $Z_{nm}^{\text{TE}} = (k_0/\beta_{nm})Z_0$  is the modal wave impedance.  $k_0$  is the free space wave number,  $\beta_{nm} = \sqrt{k_0^2 - (p'_{nm}/R)^2}$  is the propagation constant, and  $Z_0$  is the vacuum impedance. Here we consider the  $\text{TE}_{01}$  mode in a circular waveguide, since it has the lowest Ohmic loss. It is straightforward to do a similar analysis for other cavity modes.

The radiated power through an area  $S$  for the  $\text{TE}_{01}$  mode is  $P_{\text{rad}} = \frac{Z_{01}^{\text{TE}}}{2} \left( \frac{\beta_{01} p'_{01}}{R k_{c,01}^2} \right)^2 \int_S J_1^2 \left( \frac{p'_{01} r}{R} \right) d\mathbf{r}$  where  $p'_{01} = 3.83$  is the first zero of the Bessel function derivative  $J'_0(x)$  and  $k_{c,01} = p'_{01}/R$  is the wavenumber corresponding to the  $\text{TE}_{01}$  mode. The total radiation over the whole waveguide cross section is  $P_{\text{tot}} = \frac{Z_{01}^{\text{TE}}}{2} \left( \frac{\beta_{01}}{k_{c,01}^2} \right)^2 2\pi \int_0^{p'_{01}} J_1^2(x) x dx = \frac{\pi Z_{01}^{\text{TE}} \beta_{01}^2}{2 k_{c,01}^4} p_{01}'^2 J_0^2(p'_{01})$ . In the experiment, the pinhole is centered at  $r = 0$  with radius equal to  $R/5$ , so the radiation through the pinhole is

$P_{\text{pinhole}} = \frac{\pi Z_{01}^{\text{TE}} \beta_{01}^2}{k_{c,01}^4} \int_0^{p'_{01}/5} J_1^2(x) x dx$ . So the effective transmission coefficient is

$$T_{\text{exp}} = \frac{P_{\text{pinhole}}}{P_{\text{tot}}} = \frac{\int_0^{p'_{01}/5} J_1^2(x) x dx}{\int_0^{p'_{01}} J_1^2(x) x dx} \approx 0.016.$$

$T_{\text{exp}}$  is smaller than the area ratio 0.04, due to the fact that the Poynting flux of TE<sub>01</sub> mode is higher around  $r = 0.48R$ , instead of  $r = 0$ . This indicates that an off-centered pinhole would give larger transmission, which would increase the lasing efficiency (and the threshold power) by increasing the THz radiation rate compared to the Ohmic loss rate. Here, we compute the optimized transmission coefficient for a pinhole centered at  $r = 0.48R$  with the same radius  $R/5$ :

$$T_{\text{opt}} = \frac{\int_0^{2\pi} \int_0^{R_{\text{pinhole}}} J_1^2(p'_{01} x(r, \theta)/R) r dr d\theta}{2\pi \int_0^R J_1^2(p'_{01} r/R) r dr} \approx 0.075,$$

where  $x(r, \theta) = \sqrt{r_0^2 + r^2 - 2rr_0 \cos \theta}$  with  $r_0 = 0.48R$  is the pinhole center. This enhanced transmission coefficient for an off-center pinhole increases both the output power and the efficiency by decreasing the fraction of power lost to Ohmic absorption.

1. Saleh, B. E. A., Teich, M. C. & Saleh, B. E. *Fundamentals of Photonics*, vol. 22 (Wiley, 1991).
2. Panock, R. & Temkin, R. Interaction of two laser fields with a three-level molecular system. *IEEE Journal of Quantum Electronics* **13**, 425–434 (1977).
3. Everitt III, H. O. *Collisional Energy Transfer in Methyl Halides*. Ph.D. thesis, Duke University (1990).
4. Bethe, H. A. Theory of diffraction by small holes. *Physical Review* **66**, 163 (1944).
5. Bansal, R. *Fundamentals of Engineering Electromagnetics* (CRC press, 2006).