Supplementary Information (SI) for: Ultrasensitive diamond cantilever-based optical microphone

Shen Tian,[†,‡] Chaonan Lin,[†,‡] Yingying Qiao,[†] Mingyang Feng,[†] Yang Gao,[†] Kaijun Mu,[†] Mingqi Jiao,[†] Lei Li,[*,†] and Chongxin Shan[*,†]

†*School of Physics and Engineering, Zhengzhou University, China*

‡*These authous contributed equally to this work.*

E-mail: lilei@zzu.edu.cn; cxshan@zzu.edu.cn

Supporting information

- Figure S1: Optical microphone modeling.
- Figure S2: Cantilever modeling.
- Figure S3: F-P interferometer modeling.
- Figure S4: Experimental setup.
- Figure S5: Sensor design.
- Table S1: Comparison of sensitivity and MDP.

Correspondence and requests for materials should be addressed to L.L. or C.S.

S1. Performance metrics

Here, we list the definitions of relevant microphone performance metrics addressed in the main article, namely sensitivity, frequency response, signal-to-noise ratio, and minimum detectable acoustic pressure.

- **Sensitivity**: ratio between electrical output and applied acoustic pressure, typically expressed in mV/Pa or dBV for commercial microphones. For optical microphones, this overall sensitivity is a combination of electrical sensitivity, which depends on the mechanical sensitivity of the acoustic transduction element, the optical sensitivity of the optical configuration, and the system sensitivity of the optical readout system.
- **Frequency response**: sensitivity curve relative to the applied acoustic frequency, which typically has a resonant frequency peak. The general flat region of the frequency response curve is below the resonant frequency.
- **Signal-to-noise ratio (SNR)**: ratio between output in response from a reference signal (1 kHz at 1 Pa) and noise level of the microphone.
- **Minimum detectable acoustic pressure (MDP)**: acoustic detection limit of a microphone, defined as the lowest pressure for a signal-to-noise ratio S/N=1 in a resolution bandwidth (RBW) Δf.

In this work, the calculation formula for electrical sensitivity has been unified, and we provide corresponding conversion methods from different previous literature sources:[1]

$$\begin{aligned} S_e &= \frac{\Delta V}{\Delta P}(\mathrm{mV_{rms}/Pa}) = \frac{\sqrt{2}\Delta V}{\Delta P}(\mathrm{mV_{amp}/Pa}) \\ &= \frac{2\sqrt{2}\Delta V}{\Delta P}(\mathrm{mV_{pp}/Pa}) = 20\log_{10}(\frac{\Delta V/\Delta P}{\Delta V_0/\Delta P_0})(\mathrm{dBV}) \end{aligned} \tag{1}$$

where ΔV represents the root-mean-square of the output voltage, with units in RMS, amplitude, and peak-to-peak denoted as mVrms, mVamp, and mVpp, respectively. V_0 is the reference voltage (1 V) at the reference pressure P_0 (1 Pa), defined as 0 dBV. Previous works often use peak-to-peak or amplitude values of voltage for calculating sensitivity, while commercial microphones are generally calibrated with RMS values and in dBV. We have unified the units as mVamp for ease of general comparison.

S2. Transduction mechanism and modeling

A. Overall sensitivity

A typical structure of an optical microphone is shown in Fig. S1. In this work, we address the electrical sensitivity, which strongly depends on the material properties, cantilever dimensions, and optical configurations. The electrical sensitivity S_e of the diamond cantilever-based optical microphone (DCOM) can be expressed as:

$$S_e = \frac{\Delta V}{\Delta P} = \frac{\Delta P_{in}}{\Delta P} \frac{\Delta L_{cav}}{\Delta P_{in}} \frac{\Delta I_r}{\Delta L_{cav}} \frac{\Delta V}{\Delta I_r} \tag{2}$$

where V represents the output voltage of the sensing system, P represents the pressure generated by the acoustic source, P_{in} denotes the pressure difference between the two sides of the cantilever, L_{cav} represents the Fabry-Perot (F-P) cavity length, and I_r represents the reflected light intensity. Notably, in this study, the characteristics of the DCOM are tested in the near-field, so the propagation loss of the acoustic wave $\Delta P_{in}/\Delta P$ can be neglected. $\Delta L_{cav}/\Delta P_{in}$ indicates the mechanical sensitivity of the diamond material (S_m), $\Delta I_r/\Delta L_{cav}$ indicates the sensitivity (S_i) of F-P interferometry, and $\Delta V/\Delta I_r$ depends on the amplifier and the response of the photodetector in the read-out system.

B. Acousto-mechanical transduction modeling

Two typical models, the string model and the beam model, are applied to understand the dynamic mechanism of the transduction element. The string model is typically used for

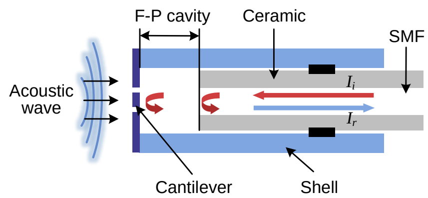

Figure S 1: Optical microphone modeling. Schematic of cantilever-based F-P optical microphone.

materials such as stainless steel, where residual stress during film processing or tensile stress applied during clamping serves as the main source of restoring force. In the case of monocrystalline diamond cantilevers, which are flexural rigid beams with rectangular cross-sections, bending stress dominates (Fig. S2). With appropriate simplifications for the thin-beam approximation and assuming no damping, the motion is determined by the Euler-Bernoulli equation:[2]

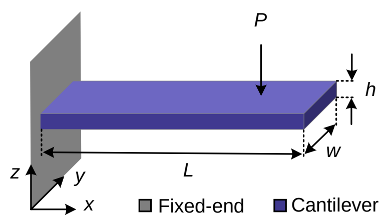

Figure S 2: Cantilever modeling. Schematic of a simple cantilever geometry.

$$F_d = \frac{\partial^2 D(x,t)}{\partial t^2}\rho S + \frac{\partial^4 D(x,t)}{\partial x^4}\widehat{E} I_x \tag{3}$$

where $D(x,t)$ denotes the displacement Δz in the z-axis direction due to the drive force (F_d), ρ denotes the mass density, and $S = wh$ denotes the cross-sectional area, where w and h represent the width and height of the cantilever, respectively. Additionally, $\widehat{E}$ denotes Young's modulus, and I_x denotes the moment of inertia of the cantilever. The solution to this differential equation describes the beam deformation as a harmonic motion with position-dependent and time-dependent terms, $D(x,t) = D_n(x)\exp(-\mathrm{i}\omega_n t)$, where ω_n represents the motion frequency and n denotes the modal number. In this paper, we address the first-order resonant frequency of the cantilever.

Considering the damping effect, the dynamic model can be simplified as a single mass model using a harmonic oscillator:

$$m_{eff}\ddot{z} + \beta\dot{z} + k_{eff}z = F\cos(\omega t) \tag{4}$$

where m_{eff} represents the effective mass of the cantilever, β is the damping constant, k_{eff} is the effective spring constant, and $F\cos(\omega t)$ denotes the sinusoidal force caused by the acoustic wave. The intrinsic damping of the cantilever represents the energy dissipation mechanism,[3] given by $\beta = \sqrt{k_{eff}m_{eff}}/Q_c$, where Q_c is the quality (Q) factor of the cantilever. The Q-factor is expressed as $Q_c = 2\pi W_0/\Delta W$, where W_0 represents the stored vibrational energy and ΔW represents the total energy dissipation per vibration cycle.

Extrinsic damping is caused by the medium and sensor geometry. For a cantilever clamped on the shell of an F-P sensor, when the cantilever oscillates, it displaces the gas molecules around it, resulting in a corresponding change in the volume of the F-P cavity.[4] The damping effect of air introduces an additional component to the effective mass and spring constant. In this case, the effective spring constant and mass of the cantilever are as follows:

$$k_{eff} = \frac{2}{3}Ew(\frac{t_c}{L})^3 + \frac{2\kappa S^2 P}{5V}, m_{eff} = 0.647m + \frac{VM}{RT}P \tag{5}$$

The solution $z(\omega)$ gives the amplitude of the cantilever as:

$$A_z(\omega) = \sqrt{z(\omega)z^*(\omega)} = \frac{F}{m_{eff}\sqrt{(\omega_0^2 - \omega^2)^2 - (\omega\beta/m_{eff})^2}} \tag{6}$$

For force sensing applications, a cantilever's sensitivity is ultimately constrained by its thermally driven motion, with a noise-limited minimum pressure P_{min}:[5,6]

$$P_{min} = \frac{P_{in}}{10^{\frac{\mathrm{SNR}}{20}} \times \sqrt{\mathrm{RBW}}} \tag{7}$$

where SNR represents the signal-to-noise ratio in decibels, and RBW is the resolution bandwidth. The comprehensive noise profile includes factors such as shot noise, thermal noise, detector and DAQ noise, etc. In this work, optical microphones were characterized under identical conditions, highlighting the predominant influence of noise originating from thermally driven cantilever vibrations. In thermal equilibrium, the mean square displacement z_{rms} is related to the spring constant through the equipartition theorem, $\frac{1}{2}k_{eff}z_{rms}^2 = \frac{1}{2}k_BT$, where T is the temperature, and k_B is the Boltzmann constant. As per Equation 5, the ultrahigh Young's modulus of the diamond cantilever suppresses thermally driven motion under temperature variation, thereby providing low thermal noise.

C. Mechanical-optical modeling

The DCOM comprises a Fabry-Perot (F-P) cavity, a 3D-printed epoxy shell, and a single-mode fiber. The F-P cavity is enclosed within the sensor shell, which is formed by two reflective surfaces: the fiber end and the inner surface of the diamond film (with a cantilever structure in the middle). The interference model can be simplified as a two-beam interferometer, as depicted in Fig. S3.

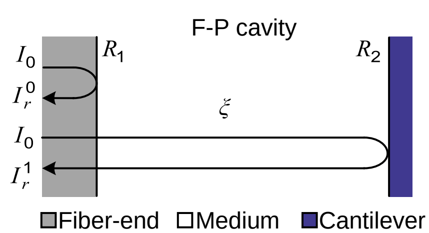

Figure S 3: F-P interferometer modeling. Schematic of F-P interference in DCOM.

The incident light can be reflected by both the diamond cantilever and the optical fiber end, and the reflective interference intensity I_r can be expressed as:

$$I_r = I_i(R_1 + \xi R_2 - 2\sqrt{\xi R_1 R_2}\cos\delta) \tag{8}$$

where I_i is the intensity of the incident light, R_1 and R_2 are the reflectivity values of the fiber end and diamond cantilever, respectively, δ is the phase difference, and ξ is the optical coupling coefficient, which is dependent on the length of the F-P cavity and the wavelength of light. ξ can be expressed as:

$$\xi = \frac{4[1+(\frac{2\lambda L_{cav}}{\pi n_0 {\omega_0}^2})^2]}{[2+(\frac{2\lambda L_{cav}}{\pi n_0 {\omega_0}^2})^2]^2} \tag{9}$$

where λ (1550 nm) denotes the wavelength of the incident laser source, L_{cav} denotes the length of the F-P cavity, here 160.19 µm for the 30 µm DCOM, n_0 denotes the refractive index of the air medium within the cavity, n_0=1, and ω_0 denotes the 1/e^2 beam radius (5.3 µm for 1550 nm). Therefore, the values of ξ are 0.015 for the 30 µm DCOM.

In a static F-P cavity, the phase difference of the two beam interference is δ=4$\pi L_{cav}/\lambda$, and the optical sensitivity S_i of the F-P interferometer is the variation in the reflected light intensity ΔI_r induced by the cavity length ΔL_{cav}, which can be expressed as:

$$S_i = \frac{\Delta I_r}{\Delta L_{cav}} = \frac{8\pi\sqrt{\xi R_1 R_2}}{\lambda} I_i \mathrm{sin}\frac{4\pi L_{cav}}{\lambda} \tag{10}$$

When L_{cav}=(2n+1)λ/8, S_i is maximized, and the operating point of the sensor is called the quadrature point (Q point). The reflected light of the optical microphone in the acoustic field can be expressed as:

$$I_r = 2I_0[1+\gamma\mathrm{cos}(\frac{4\pi(L_{cav}+\Delta L)}{\lambda}+\pi)] \tag{11}$$

D. Optical-electrical modeling

The system sensitivity is mainly determined by the response of the photodiode. The response of balanced PD (PDB450C, Thorlabs) can be expressed as:

$$\Delta V = \Delta I_r \times \mathrm{R}(\lambda) \times \mathrm{G} \tag{12}$$

where R(λ) denotes the responsivity of the photo diode at a given wavelength and G denotes the transimpedance gain of the PD output. The sensor is connected to the sensing system, and the relation between the defelction of the diamond cantilever (Δz) and the output voltage of the photodetector (ΔV_r) can be expressed by the following equation:

$$\Delta z = \frac{\lambda(R_1 + \xi R_2)}{8\pi V_0 \sqrt{\xi R_1 R_2}} \Delta V_r \tag{13}$$

where V_0 ($V_0 = 4.9$ mV) is the output voltage from the PD at the Q point.

S3. Additional data on the measurement of F-P interference

The experimental setup of F-P interference is shown in Fig. S4 a. Incident light from a SLED source (1550 ± 30 nm) was transmitted through a circulator, reflected by the sensor head, and detected by a fiber Bragg grating interrogation analyzer (FBGA, Bayspec). The FBGA has a frequency response time of 5 kHz and is connected to the PC via a USB interface. The use of FBGA provides a real-time and high-speed measurement of the F-P interference spectrum.

Fringe visibility, defined by $V = (I_{max} - I_{min})/(I_{max} + I_{min})$, represents the optical interference within the optical microphone, where I_{max} and I_{min} are the maximum and minimum intensities of reflected light, respectively. Following the assembly of the cantilever and shell, all fabricated sensors mentioned in the main article underwent testing for fringe visibility. The ceramic ferrule in the shell was adjusted to reach maximum fringe visibility. Figures S4 b and S4 c represent the interference spectra of the 30 µm (red lines) and the 50 µm (blue lines) cantilevers, respectively. As depicted in Fig. S4 b, the spectral combs of the 30 µm and 50 µm DCOMs exhibit similar fringe visibility (approximately 6 dB) near 1540 nm. Considering the use of a 1550 nm DFB laser in the sensor characterization section, we

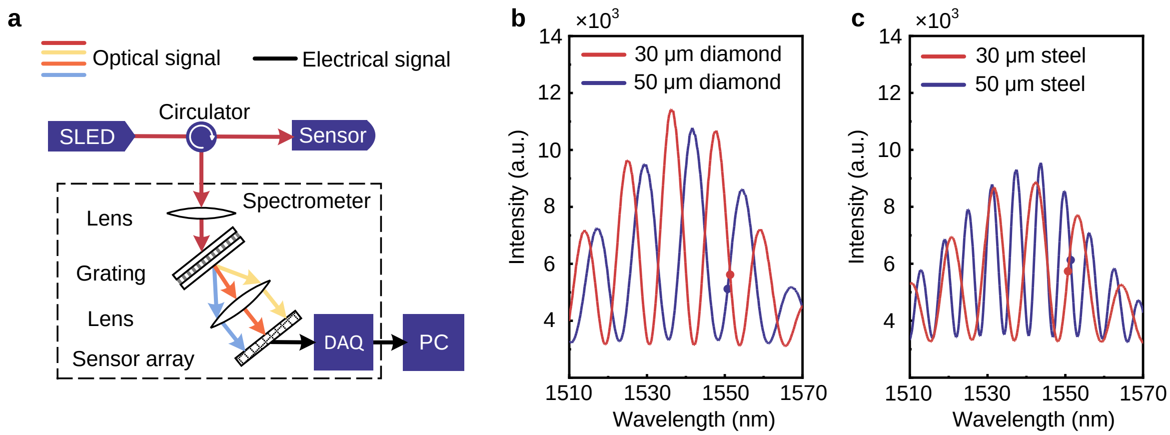

Figure S 4: Measurement of F-P interference. **a**, Experimental setup. **b**, Interference spectra of the proposed DCOMs. The Q-points of 30 µm (1551.4 nm, red dot) and 50 µm (1550.7 nm, blue dot) cantilevers are marked on curves. **c**, Interference spectra of sensors using stainless steel cantilevers.

estimated the Q points Q_1 (1551.4 nm) and Q_2 (1550.7 nm) near 1550 nm, as depicted in Fig. S4 b and Fig. 1d of the main article. Moreover, in Fig. S4 c, sensors using stainless steel cantilevers exhibit similar fringe visibility of 5 dB, with the estimated Q points Q_3 (1550.5 nm) and Q_4 (1550.8 nm).

In the sensor characterization experiment, the wavelength of the DFB laser (1550 $\pm$ 1 nm, Fig. 3a of the main article) was adjusted to match these Q points to maximize the optical sensitivity S_i of the sensors.

S4. Additional data on comparison group

We fabricated two optical microphones with 30 µm and 50 µm stainless steel cantilevers as a comparison group. The sensor structure, shown in Fig. S5 a, is identical to the proposed DCOMs (Fig. 2g in the main article). The stainless steel membrane with a cantilever structure (Fig. S5 b) is secured by two clamps and enclosed in the shell. The assembled sensor is depicted in Fig. S5 c.

The performance of the stainless steel sensors was tested under identical procedures and conditions after the DCOMs were tested. We demonstrated the output signal as a function

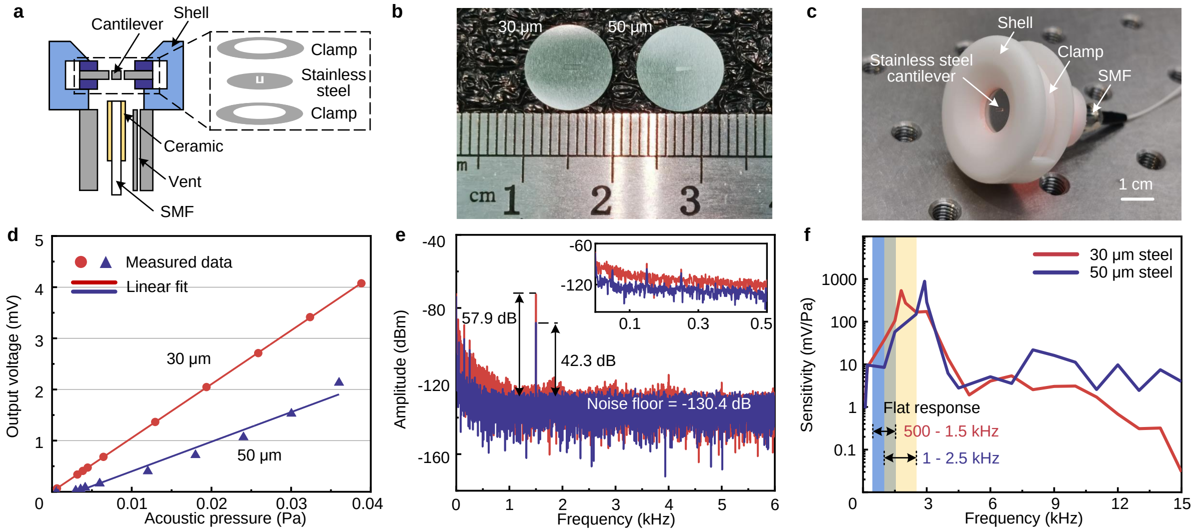

Figure S 5: Sensor design and characterization. **a**, Schematic of the assembly structure. **b**, Photograph of micromachined 30 µm and 50 µm stainless steel membranes with cantilever structures. Cantilever length, 3 mm. Cantilever width, 600 µm. **c**, Illustration of the assembly sensor. **d**, Linear fitting of output voltages as a function of applied pressure for 30 µm (red dotted line) and 50 µm (blue triangle line) sensors. Acoustic frequency, 1.5 kHz. **e**, SNR spectra of **d** under an acoustic pressure of 5 mPa. The inset shows a zoomed-in low-frequency noise ranging from 0 - 0.5 kHz. **f**, Frequency response.

of the 1.5 kHz applied acoustic pressure ranging from 0.5 to 30 mPa (Fig. S5 d). The 30 µm sensor achieved a sensitivity level of 105.0 mV_{amp}/Pa, surpassing the 57.8 mV_{amp}/Pa observed for the 50 µm sensor. Notably, the fitting curves of the 50 µm sensor exhibited limited linearity (R^2=0.987), especially under pressures less than 10 mPa, demonstrating the limited sensitivity of the 50 µm stainless steel cantilever for weak acoustic signal sensing. Figure S5 e displays the SNR spectra detected by the 30 µm (red line) and 50 µm (blue line) sensors under 5 mPa acoustic pressure. The SNR of 57.9 dB is observed at a frequency of 1.5 kHz, which corresponds to the MDP of 6.4 $\mu\text{Pa}/\sqrt{\text{Hz}}$ at a 1 Hz resolution bandwidth. The frequency response of the sensors is shown in Fig. S5f and Fig's 3d and 3e of the main article. In Fig. S5f, the 30 µm sensor demonstrated a response bandwidth spanning from 500 Hz to 1.5 kHz (blue region), with a resonant frequency of approximately 1.8 kHz. Additionally, the 50 µm sensor exhibited a response from 1 kHz to 2.5 kHz (yellow region), with a resonant frequency of approximately 2.9 kHz. When compared to the DCOMs, the stainless steel sensors show significantly lower sensitivity and a narrower response bandwidth.

S5. Sensitivity and MDP comparison

The performance metrics compared in the main article are listed in Table S1.

References

(1) Kejing Ma et al. A wave-confining metasphere beamforming acoustic sensor for superior human-machine voice interaction. *Sci. Adv.* **8**, eadc9230 (2022) .

(2) Boisen A, Dohn S, Keller S S, et al. Cantilever-like micromechanical sensors. *Rep. Prog. Phys.* **74**(3): 036101 (2011) .

(3) Héritier M, Eichler A, Pan Y, et al. Nanoladder cantilevers made from diamond and silicon. *Nano Lett.* **18**(3): 1814-1818 (2018) .

(4) Koskinen V, Fonsen J, Roth K, et al. Progress in cantilever enhanced photoacoustic spectroscopy. *Vib. Spectrosc.* **48**(1): 16-21 (2008) .

(5) Bucaro J A, Lagakos N, Houston B H, et al. Miniature, high performance, low-cost fiber optic microphone. *J. Acoust. Soc. Am.* **118**(3): 1406-1413 (2005) .

(6) Wang Y, Rong J, Xing E, et al. A Lever-Type PDMS Flexible Cavity for Acoustic Vector Sensor With High Sensitivity[J]. *IEEE Sens. J.* **23**(6): 5637-5642 (2023) .

Table S1: Sensor comparison. Comparison of the sensitivity (mV_{amp}/Pa) and MDP (µPa/$\sqrt{Hz}$) of the 30 µm DCOM with existing microphones using various signal transductions.

Journal [Ref. in main article] or commercial mic. (Company)	**Signal transduction**	**Materials/ structures**	**Sensitivity**	**MDP**
Adv. Mater. **34**, 2109545 (2022) [8]	Condenser	Polymer	31.7	NA
ACS Appl. Mater. Interfaces. **9**, 1237-1246 (2017) [25]	Condenser	Graphene	141	100
J. Microelectromech. S. **24**, 241-248 (2015) [10]	Condenser	Silicon	300*	6
Type 4966 (B&K, Denmark)	Condenser	Stainless steel	67.2	8
Sci. Adv. **7**, eabe5683 (2021) [6]	Piezoelectric	PZT membrane	73.5*	NA
Nat. Commun. **7**:11108 (2016) [9]	Piezoelectric	Polymer/Nanofiber web	376.2*	NA
Sci. Adv. **3**, eaas8772 (2018) [2]	Triboelectric	Silver nanowire	49.5*	NA
Sci. Robot. **3**, eaat2516 (2018) [5]	Triboelectric	Polymer	369.1	NA
IEEE Sens. J. **21**, 17882-17888 (2021) [21]	Optical	Gold membrane	17.8*	95.3
Microelectron. Eng. **166**, 50-54 (2016) [23]	Optical	Silver membrane	820.24	83
J. Lightw. Technol. **36**, 5650–5655 (2018) [24]	Optical	Polymer membrane	342.78*	17.9
IEEE Photonic. Tech. L. **25**, 932-935 (2013) [26]	Optical	Graphene membrane	18.6	60
J. Lightw. Technol. **36**, 5224–5229 (2018) [19]	Optical	Corrugated MEMS membrane	91.78*	2.97
IEEE Trans. Instrum. Meas. **67**, 1994-2000 (2018) [22]	Optical	Corrugated silver membrane	1414*	86.97
Opt. Express **29**, 16447-16454 (2021) [18]	Optical	Glass/No membrane	282.3*	530
Opt. Lett. **45**, 3516-3519 (2020) [15]	Optical	Polymer/Spiral	167.3	0.33
Nano Lett. **19**, 2627-2633 (2019) [14]	Optical	Biomimetic	15*	1665
Microsyst. Nanoeng. **9**, 65 (2023) [16]	Optical	CaF_2 microring	11540*	9.4
Opt. Express **31**, 21796-21805 (2023) [13]	Optical	Steel cantilever	1058*	6.2
Sens. Actuators, A **279**, 107-112 (2018) [27]	Optical	Steel cantilever	174.6	8.5
This work	Optical	Diamond cantilever	1433.79*	0.24

* Sensitivity values were reported at resonance frequencies