Bayesian Optimization for Practical H_2 Sensors: Inverse Design of Pd-based Plasmonic Metasurfaces

Pernilla Ekborg-Tanner,[1] Athanasios Theodoridis,[1] Joachim Fritzsche,[1]
Christoph Langhammer,[1] Andrea Baldi,[2] and Paul Erhart[1,3]

[1] *Department of Physics and Astronomy, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*
[2] *Department of Physics and Astronomy, Vrije Universiteit Amsterdam, De Boelelaan 1100, 1081 HZ Amsterdam, Netherlands*
[3] *Wallenberg Initiative Materials Science for Sustainability, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden**

Hydrogen detection is becoming increasingly important as its use grows across energy and industrial systems. Optical sensing platforms based on palladium (Pd) nanoparticles are attractive for this task because hydrogen uptake directly alters their plasmonic response. Organizing such nanoparticles into periodic two-dimensional arrays, known as metasurfaces, further enhances their optical response through collective resonances. However, the large design space presented by chemical composition, nanoparticle geometry, and array structure calls for systematic approaches for optimizing complex nanoalloy metasurface geometries. Here, we develop an inverse-design framework based on Bayesian optimization that couples first-principles dielectric functions with electromagnetic simulations to identify high-performance PdAu nanodisk arrays for hydrogen sensing in the 1–100 mbar range where the flammability of H_2 becomes a concern. We use our approach to search a five-dimensional design space, comprising nanodisk height and radius, array pitch, polymer coating thickness, and Au fraction in order to maximize the H-induced change in extinction at a single wavelength of choice. The results show that integrating first-principles optical models with data-efficient optimization yields experimentally feasible nanoparticle metasurfaces tailored for targeted hydrogen pressures, while providing a pathway to future multiobjective sensor design. They also reveal remaining gaps in the modeling methodologies that still limit the quantitative reliability of the approach.

In the search for sustainable alternatives to fossil fuels, hydrogen gas (H_2) has emerged as a promising energy carrier. To achieve a safe hydrogen economy, efficient and reliable H_2 sensors are essential due to the high flammability of H_2 under ambient conditions. Over the past decades, nanostructured palladium (Pd)-based materials have shown great promise for this purpose, providing a spark-free, optical sensing platform [1–10]. This functionality is enabled by the ability of Pd to quickly absorb and release hydrogen, which simultaneously changes its optical properties. In Pd nanodisks, this effect manifests as a shift of the localized surface plasmon resonance upon H_2 exposure, which can be utilized as a sensing mechanism [1].

Pd is, however, associated with several challenges in the context of efficient H_2 sensing [11], including hysteresis during H uptake and release [12, 13], CO poisoning [14], slow kinetics [15], low-quality localized surface plasmon resonances [16–18], and poor sensitivity in humid environments [19]. These challenges have motivated extensive research over the past decade, resulting in mitigation strategies such as alloying [5, 14, 19], applying coatings [7, 20, 21], nanostructuring [18], and machine learning-based signal processing [15]. With these efforts, Pd-based nanomaterials have evolved into a highly tunable platform for plasmonic H sensing. Current research efforts are, however, typically limited to trial-and-error approaches targeting each challenge in isolation, and a unifying perspective that systematically connects the thermodynamics, kinetics, optical response, and practical device constraints remains absent.

To this end, we propose here a flexible computational framework to model and optimize sensor systems. We focus on sensors based on PdAu nanodisks due to the hysteresis reduction and improved kinetics during H sorption associated with Au [13, 22]. The nanodisks are further organized into ordered 2D arrays, or metasurfaces, leading to narrower optical resonances, known as surface lattice resonances (SLRs), arising from their collective optical response [18, 23]. This phenomenon requires a uniform refractive index surrounding the nanodisks, achieved by coating the arrays with PMMA, which matches the refractive index of the silica substrate while having a limited impact on H_2 absorption rates in the nanodisks [7]. The optical response of the alloy material is described by its macroscopic dielectric function (DF), obtained from first-principles density functional theory calculations, as a function of the Au and H concentrations. The sensor system is characterized based on continuum-scale electrodynamic simulations of the extinction spectrum before and after hydrogenation.

This framework is deployed in Bayesian optimization (BO) to optimize sensor design with respect to the five-dimensional parameter space spanned by nanodisk height and radius, array pitch (disk center-to-center distance), PMMA layer thickness, and Au concentration in the alloy. BO is an iterative process where a surrogate model is used to predict an objective function in the entire domain while quantifying the associated uncertainty. This guides the selection of points to sample by balancing exploration of unsampled regions and exploitation of regions close to optima.

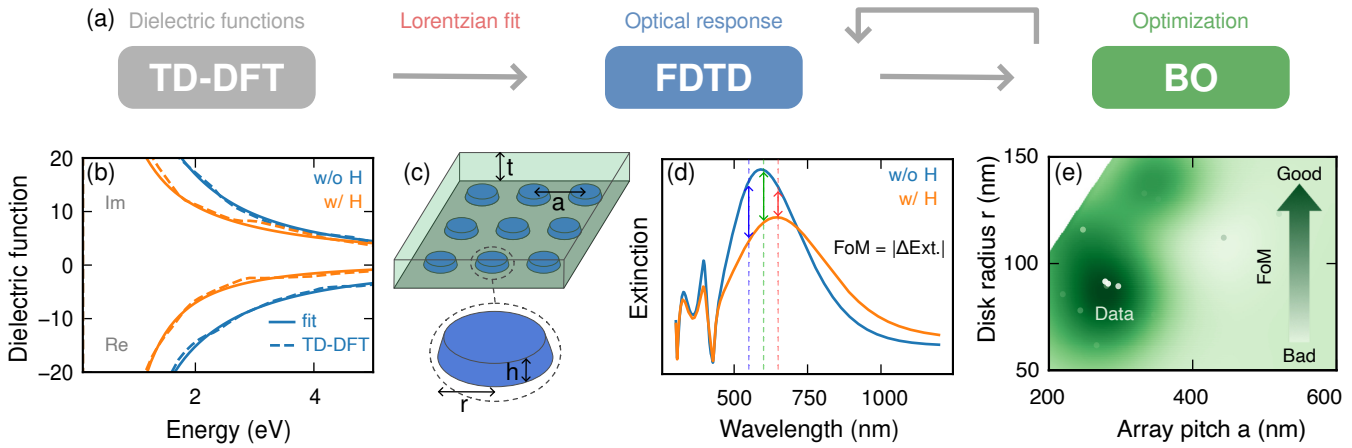

FIG. 1. **Overview of the optimization procedure.** (a) Flowchart of the computational methods involved. (b) Calculated and fitted dielectric function before and after hydrogenation used to simulate the extinction spectra. (c) Schematic illustration of the sample geometry including the involved design parameters corresponding to disk radius r and height h, array pitch a, PMMA layer thickness t, and Au concentration in the PdAu alloy. (d) Simulated extinction spectra for a nanodisk array before and after hydrogenation, used to calculate the sensing FoM from the extinction shift at a selected wavelength. (e) The predicted FoM landscape from the BO surrogate model, used to iteratively suggest new geometries.

In the present work, the objective function is the sensor figure of merit (FoM), which can be defined as any function of the extinction spectrum at different levels of hydrogenation. We explore different FoMs but mainly focus on a cost-efficient single-wavelength sensing principle by defining the FoM as the change in extinction at a selected wavelength upon hydrogenation [24–26]. Compared to typical academic approaches, where the full extinction spectrum is recorded, this represents a reduced set of observables, lowering the cost and complexity of the required setup and helping to bridge the gap between academia and industry, since single-wavelength readout is typically preferred for real sensor devices.

Furthermore, our approach takes into account the complicated thermodynamics of H absorption by explicitly including the non-linear dependence of the absorbed H concentration on the surrounding H_2 partial pressure. This is achieved by optimizing sensors for a specific target partial pressure and avoiding the Au concentration-dependent hysteresis region, enabling high sensitivity across the relevant partial pressure interval. In the present work, we focus on the 1–100 mbar range, which covers the detection limits set by the U.S. Department of Energy corresponding to 1 mbar and 40 mbar for automotive and stationary applications, respectively [27], with the latter coinciding with the lower flammability limit of H_2 in air. Kinetics and environment-dependent factors, such as CO or the role of humidity, are not explicitly considered in the present work, but the framework could be extended to multiobjective optimization and even multiplexed sensors in the future.

In the following, we first describe the computational framework, which combines first-principles DF calculations with electrodynamic simulations and BO for sensor design. We then present the optimization results for single-wavelength sensors targeting H_2 partial pressures in the 1–100 mbar range, and show that this approach efficiently identifies experimentally feasible designs with good sensing performance. This is followed by experimental validation of the predicted structures, where we find good overall agreement with simulations despite remaining discrepancies. We attribute remaining deviations primarily to the accuracy of the calculated DF, and discuss how these could be systematically improved in future work. We conclude by discussing the implications of our findings for practical sensor design and future extensions of the optimization framework.

RESULTS AND DISCUSSION

The computational framework (Fig. 1a) starts at the density functional theory-level with the set of DFs for the bulk PdAuH system calculated in Ref. 17, which are fitted to a Lorentzian representation (see Supplementary Note 1) to allow for smooth and continuous changes with Au and H concentrations (Fig. 1b, Supplementary Note 1). Each candidate sensor consists of a square periodic array of nanodisks on a silica substrate coated with PMMA (Fig. 1c). The nanodisks have tapered sides, with an angle of 60° between the sidewall and the base plane, reflecting the truncated cone geometry that results from physical vapor deposition through a lithographic mask. The sensor geometry is defined by five parameters: nanodisk radius r (50–150 nm), height h (20–80 nm), array pitch a (200–600 nm), PMMA thickness t (0–300 nm),

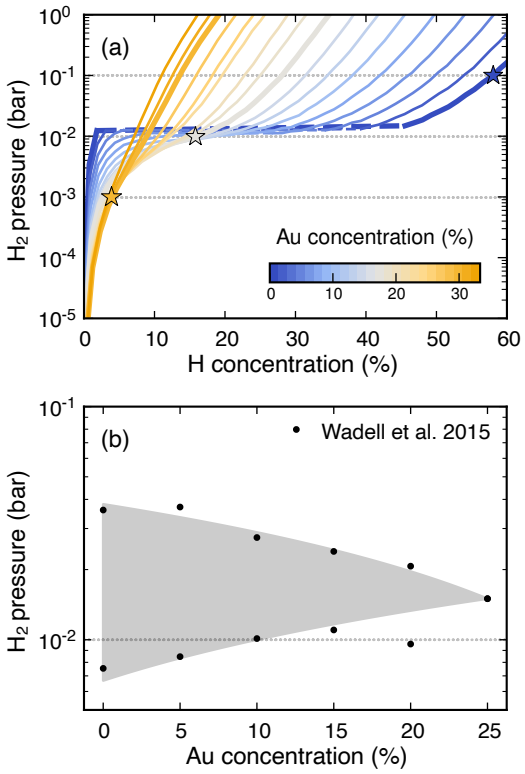

FIG. 2. **Thermodynamic model.** (a) H uptake pressure-composition isotherms for different alloy compositions at 300 K from Ref. 28. The stars indicate the compositions of the optimized 650 nm single-wavelength sensors at each target pressure. (b) Hysteresis gap as a function of alloy composition based on Ref. 13 (dots) and the corresponding linear fit of the gap region limits (limits of the shaded region).

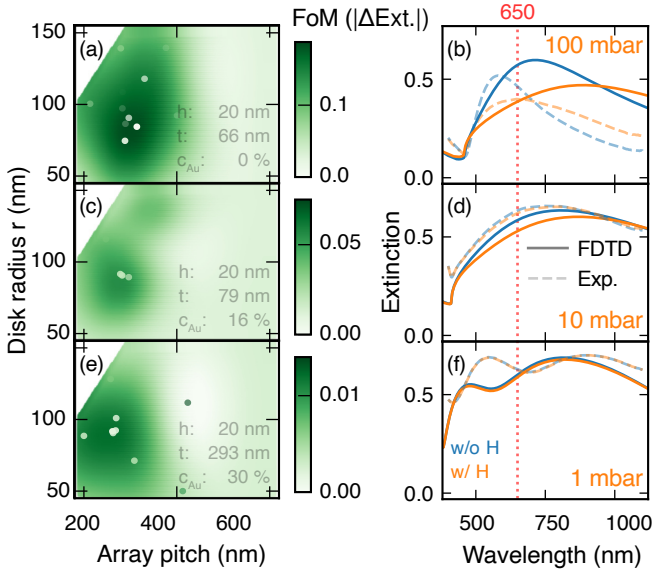

FIG. 3. **Optimization of the @650 nm single-wavelength sensors.** Two-dimensional slices of the final predicted FoM landscape for optimization at target pressures 100 (a), 10 (c) and 1 mbar (e), with the values of the other design parameters annotated. The corresponding extinction spectra (b, d, f) from simulations, as well as from experimental measurements, on the fabricated samples.

and Au concentration c_{Au} (0–30 %) in the PdAu alloy.

The extinction spectra of each configuration before and after hydrogenation to the target H_2 partial pressure (this mapping is explained below) are obtained from finite-difference time-domain (FDTD) simulations, and the sensor performance is quantified through the selected FoM (Fig. 1d). Here, we focus on a single-wavelength FoM defined as the absolute change in extinction at a selected wavelength, and specifically 450, 550, or 650 nm, chosen to represent wavelengths of commercially available blue, green, and red lasers. We also consider an alternative full-spectrum FoM defined as the shift in the plasmon peak upon hydrogenation divided by the full width at half maximum (measured before hydrogenation).

BO [29] is employed to efficiently explore the resulting five-dimensional design space with the goal of maximizing the FoM. Here, a Gaussian process surrogate model predicts the FoM landscape based on the available data, starting from an initial set of randomly selected data points (Fig. 1e). In each iteration, new candidate configurations are selected by maximizing an acquisition function that balances exploitation of current knowledge about the optimum with exploration of regions of high model uncertainty (Fig. S1). In practice, the optimization is run for 30 iterations, after which the FoM typically shows only marginal further improvement (Fig. S1c). This is followed by five iterations of pure exploitation to ensure that the predicted maximum is verified by simulations. The best-performing configuration is then selected for experimental validation.

The H concentration of the hydrogenated state, c_{H}, is determined from thermodynamic data from Ref. 28, which relates the target H_2 partial pressure to the equilibrium absorbed bulk H concentration of the nanodisks (Fig. 2a). The hysteresis region as a function of Au concentration (shaded region in Fig. 2b) is estimated based on data from Ref. 13 and excluded during the optimization, effectively setting a target pressure-dependent lower limit for the Au concentration. For instance, at 10 mbar H_2, the Au concentration needs to be at least 10 %, as lower concentrations are associated with hysteresis at this pressure. In addition to spanning established detection targets [27], the chosen target pressures of 1, 10, and 100 mbar represent distinct thermodynamic regimes on both sides of the plateau pressure with and without hysteresis, enabling benchmarking across the practically relevant pressure range.

Single-wavelength sensor @650 nm

To illustrate the optimization procedure, we first present results for sensors optimized for extinction shift at 650 nm and target pressures of 1, 10, and 100 mbar, noting that similar results are obtained at 450 and 550 nm, as discussed in the following section. The final predicted FoM landscapes show a single broad maximum for all target pressures (Fig. 3a, c, e). The locations of the maxima are similar across target pressures, except for the Au concentration (see annotations in Fig. 3a, c, e), which increases as the target pressure decreases; this is expected since the H absorption increases with Au concentration below the plateau pressure (Fig. 2a). The PMMA thickness varies between the optimized samples, but this is due to an overall weak effect of the PMMA thickness on performance rather than a strong preference (Fig. S1).

The design parameters at the predicted optima are rounded to facilitate fabrication (Table S1) and the resulting sensor geometries are fabricated and their response to pulses of H_2 is measured. The overall spectral shapes show qualitative agreement between experiment and simulation, though systematic shifts of the peaks are observed in the measurements (Fig. 3b, d, f). These deviations can be attributed to two main factors. First, the fabrication process is associated with a noticeable uncertainty in the final size parameters (Fig. S2), in particular due to the thermal annealing necessary to induce alloy formation [30], which commonly reduces the aspect ratio and causes blueshifts of the plasmon peak [17]. Second, the first-principles-based DFs of the plasmonic material are subject to the inaccuracies associated with the underlying exchange-correlation functional in the density functional theory calculations [31], as well as the phenomenological treatment of the Drude peak [32], which is fundamentally sample-dependent and would differ between an idealized material and a sample with defects. In addition, the magnitude of the H-induced optical shift is smaller in the experiments than in the simulations (see, for instance, Fig. 3d), which alludes to an overestimation of the H absorption in the thermodynamic model.

To fully validate the performance of the fabricated sensors, each sample was exposed to a sequence of H_2 pulses with partial pressures ranging from 0.6–100 mbar (Fig. 4a). We find that, as intended, the sensors have a clear and reversible response at their target pressure. The 10 mbar sensor performs well over the entire pressure range, whereas the 100 mbar sensor has almost no signal in the lower pressure interval. The 1 mbar sensor, while having an overall clear signal across the pressure range, shows worse kinetics and a lower amplitude optical readout compared to the other sensors. These results confirm that the optimized sensors realize high sensitivity across a broad range of pressures down to 0.6 mbar H_2 partial pressure.

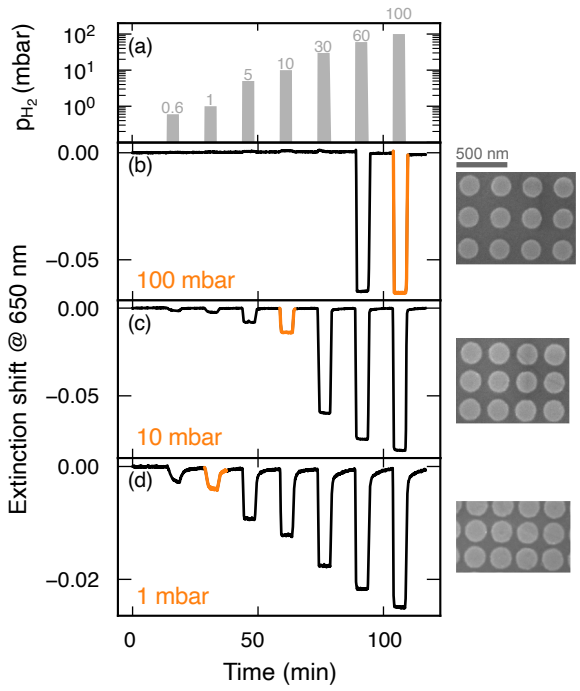

FIG. 4. **Time-resolved optical response of the @650 nm single-wavelength sensors.** (a) Sequence of hydrogen pulses ranging from 0.6–100 mbar used to test the optical extinction change at a wavelength of 650 nm for the sensors optimized for the (b) 100, (c) 10, and (d) 1 mbar target pressures. The insets to the right show scanning electron microscope images of the samples.

All single-wavelength sensors compared

Sensors were also optimized for single-wavelength shifts at 450 and 550 nm with similar results (Fig. S3). The predicted FoM landscapes are similar across all wavelengths, as well as target pressures, with a slight shift of the global maxima toward smaller array pitch and disk radius with decreasing wavelength, consistent with the size-dependent blueshift of the plasmonic resonance. The weak dependence on the PMMA thickness remains for all single-wavelength sensors, with the optimized PMMA thickness spanning the full investigated range without a systematic trend. The primary function of the PMMA layer is to enable the formation of SLRs. However, the extinction spectra of the best-performing samples exhibit broad peaks near the operational wavelength, characteristic of localized surface plasmon resonance behavior rather than the narrow features typically associated with SLRs. This observation explains the limited impact of the PMMA thickness on the optimization results. All optimal configurations feature densely packed arrays, which can be attributed to the higher effective plasmonic surface area enhancing the absolute extinction and, consequently, the magnitude of the H_2-induced optical shift. While the initial motivation for

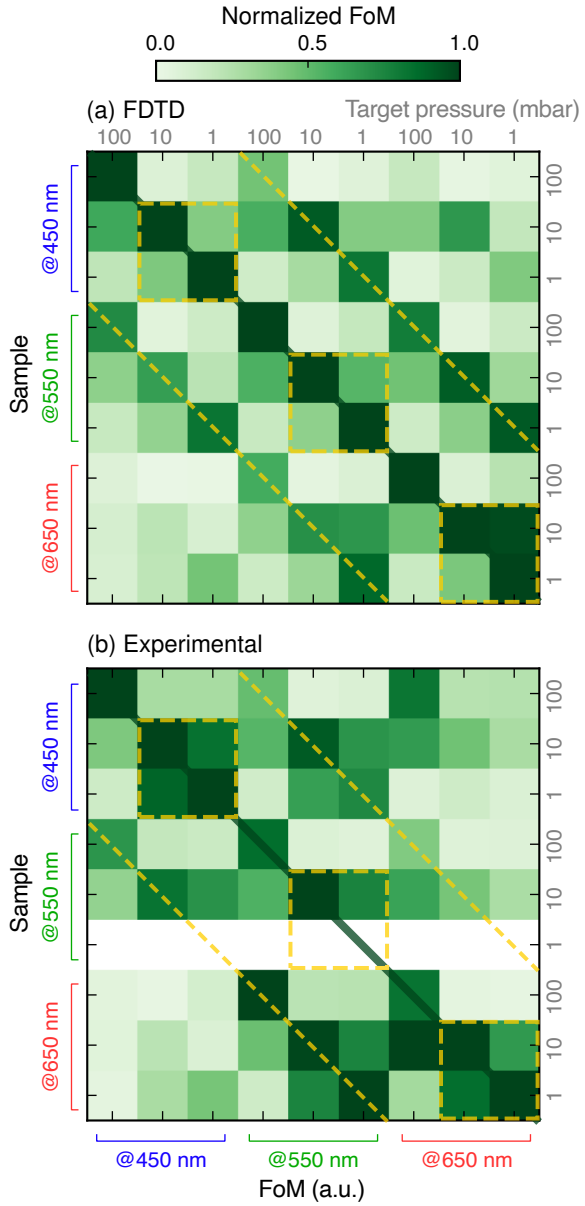

FIG. 5. **FoM comparison for all single-wavelength sensors.** All considered FoMs for all samples from the simulations (a) and experimental verification (b). The lower x-axis indicates the FoM and the upper x-axis further specifies at which target pressure the FoM is evaluated. The y-axis corresponds to the samples ordered such that their respective target FoM follows the same order as the x-axis. The darkest color indicates the highest FoM in each column, which is expected to coincide with the dark green diagonal, while the yellow dashed lines highlight secondary observed features. Note that the sample corresponding to the sixth row of the lower panel is missing due to fabrication issues.

using ordered rather than random arrays was that the resulting SLRs would provide narrower optical features and thus improved sensing performance, these results suggest that, for these FoMs, the main advantage actually arises from the higher achievable packing density.

In Fig. 5a we plot the normalized theoretical FoMs of all single-wavelength sensors investigated in this study. We find a significant increase in performance across the diagonal, indicating the success of our approach at tailoring sensor performance to specific pressure ranges. In addition, a pattern of additional diagonal lines (yellow dashed lines) shifted by three steps appears. This feature corresponds to a large extinction shift at a wavelength ±100 nm from the targeted wavelength, at the same target pressure, which is expected due to the broad plasmon peaks of these systems. Similarly, there is a correlation between the sensors optimized for the two lower target pressures (yellow dashed squares), where sensors optimized for 10 mbar typically perform well for 1 mbar and vice versa. This correlation likely emerges because, at the Au concentrations considered here, both pressures fall in similar thermodynamic regimes and thus exhibit similar H_2 absorption behavior (Fig. 2a).

The same comparison for the experimentally fabricated and measured samples shows mostly similar behavior (Fig. 5b) with six out of the eight fabricated samples being the best performing sensor at their intended target. Due to the aforementioned discrepancy between the simulations and experimental measurements, some deviations from the simulation-based results are evident. Most notably, the off-diagonal to the left (yellow dashed line in Fig. 5b) is very prominent for several samples, indicating good performance at a wavelength 100 nm lower than the target. This is consistent with two effects observed in many of the fabricated samples relative to the simulations: a blueshift of the main plasmon peak and an increase in relative amplitude of peaks at shorter wavelengths with respect to those at longer wavelengths (Fig. S3b). Both of these effects yield a larger extinction at the 100 nm lower wavelength, leading to an increased FoM. In addition, the correlation between the sensors at the lower target pressures (yellow dashed squares) predicted theoretically is also visible in the measured response.

Lastly, we have also optimized sensors geometries for the same target pressures but for a FoM based on the full spectral readout, corresponding to the H-induced peak shift divided by the full width at half maximum (Fig. S5). This approach shows a large theoretical improvement in performance, but the global maxima are much more narrow compared to the single-wavelength sensors, making the experimental verification more difficult. As a result, the full-spectrum sensors display a poor signal-to-noise ratio (Fig. S6) and are unexpectedly outperformed by one of the single-wavelength sensors (Fig. S7). This highlights that there is considerable potential for future devel-

opment if the agreement between experiment and modeling can be improved further.

CONCLUSIONS

We have established a BO framework for the inverse design of PdAu metasurface H_2 sensors that integrates optical simulations with first-principles-based DFs for PdAu alloys in the pristine and hydrogenated states. Applied to sensors targeting defined H_2 pressures, it successfully identifies high-performance regions in the design space. These devices meet the U.S. Department of Energy [27] requirements for the limit of detection using only a single-wavelength readout, demonstrating that computationally guided design can translate directly to practical, manufacturable sensors. This approach could be extended to design multiplexed sensors, where different regions of the sensor surface target different pressure regions, and multiobjective optimization where, for instance, response kinetics, gas selectivity, or stability in humid conditions, are treated simultaneously, provided that the corresponding physical phenomena can be modeled correctly.

While the agreement between simulations and experiments is good given the first-principles basis of the DFs without any empirical adjustment, this approach is limited by the remaining discrepancies which becomes critical near narrow optima. To reach the full potential of this approach, the sources of disagreement need to be reconciled, including (i) improved control over the geometry in the fabrication process, (ii) refinement of the DFs of the nanodisks to represent real materials, (iii) adjustment of the thermodynamic model for accurate H absorption isotherms, and (iv) refinement of the optical simulations to better describe the entire optical system. In particular improved structural control and refined material models will be essential to fully exploit the predictive capability of this optimization framework, and will therefore have to be targeted in future work. Once these limitations are overcome, optimization frameworks of this kind will eliminate tedious trial-and-error and grid searches associated with traditional sensor development and enable efficient and rapid design across increasingly complex parameter spaces and objective functions.

METHODS

Bayesian optimization

We use BO [29] to iteratively optimize the FoM as a function of nanodisk height and radius, array pitch and PMMA thickness. We consider different FoMs, which are always calculated from extinction spectra with and without H in relation to the target pressure *via* FDTD simulations. The optimization relies on an underlying Gaussian process which predicts the mean and variance of the FoM landscape in each step. We use the open source python package GPy [33] to set up Gaussian processes. The covariance function consists of radial basis function kernels, one for each optimization parameter (*i.e.,* a total of four), with Γ functions as hyperpriors for the length scale and variance of the kernels.

The FDTD simulations are the by far the most time consuming step. In order to maximize efficiency, two sensor geometries are suggested in each iteration and run in parallel. The geometries are suggested based on maximizing an acquisition function that balances exploration and exploitation. We use the separable natural evolution strategy [34] as a global optimization algorithm to maximize an upper confidence bound acquisition function [29], with a constraint that ensures that the second geometry picked in every iteration is not within same hyper-rectangle surrounding the first (there is no such restriction between iterations).

Lorentzian fit of DFs

The DFs used in this work are based on the set of $Pd_{1-x}Au_xH_y$ DFs calculated in Ref. 17 with time-dependent density functional theory. The original calculations are performed with 24 (metal) atom unit cells which limits the composition resolution to $1/24 \approx 4\,\%$. In order to allow for continuous changes of the DF with composition, the DF is fitted to a Lorentzian representation parametrized by polynomials of the compositions which results in smoother DFs, seemingly in better agreement with previous experimental work (Fig. S4). Details of this procedure can be found in Supplementary Note 1.

FDTD simulations

Extinction spectra were calculated using the FDTD method implemented in Ansys Lumerical FDTD [35]. The computational cell consists of a truncated cone $Pd_{1-x}Au_xH_y$ nanodisk placed on a silica substrate with refractive index 1.46. The nanodisk is embedded in a PMMA layer with refractive index of 1.51 based on experimental measurements [36]. To simulate an ordered array, periodic boundary conditions are applied in the x and y-directions while a so-called perfectly matched layer is used in the z-direction. The sample is illuminated by a plane wave source in the wavelength interval 250–1250 nm at normal incidennce. The hydrogen induced lattice expansion is neglected to avoid optical shifts caused by how the nanodisk is represented on the grid.

Sample fabrication

Nanodisk arrays were fabricated on 100 mm fused silica wafers (Corning 7980, 500(20) µm thickness, double-side polished, Ra < 1 nm). A bilayer electron beam lithography resist stack was prepared by spin-coating AR-P 6200 (1:2 in anisole, 4000 rpm, 5 min bake at 180 °C) followed by PMMA 950k A2 (4000 rpm, identical bake). A 20 nm Cr discharge/reflective layer was thermally evaporated (Lesker NANO36) prior to exposure.

Patterning was performed by electron beam lithography (Raith EBPG 5200; 10 nm beam step size, 20 nA). Proximity-effect correction was applied using a Monte Carlo–based dose modulation routine. After exposure, the Cr layer was removed by a 40 s Cr wet etch.

The resist stack was developed in two steps: 1 min in H_2O:isopropyl alcohol (1:4) for PMMA development, followed by 1 min in o-xylene for AR-P 6200. A brief O_2 plasma descum (5 s, 50 W, 0.333 mbar (=250 mTorr), 80 sccm O_2; Plasma-Therm Batchtop) was carried out prior to metallization. Metal deposition was performed in a Lesker PVD 225, followed by overnight lift-off in Remover 1165 and sequential rinsing (acetone, isopropyl alcohol, de-ionized water).

For dicing, samples were protected with S1813 (2000 rpm, 1 min bake at 110 °C) and diced using a DISCO DAD3350 with a resin-bonded diamond blade (K010-600JXS; 2 mm/s feed, 25 000 rpm). The protective coating was removed by two acetone baths (5 min each) and subsequent isopropyl alcohol/de-ionized water rinsing.

Experimental measurements

The setup used for evaluating the performance of the samples consists of a quartz tube plug-flow reactor with optical access for transmittance measurements (Insplorion AB), and is equipped with a set of mass flow controllers (Bronkhorst High-Tech B.V.) that control the flow rate and gas composition (argon, H_2) at atmospheric pressure. The reactor temperature was controlled using a closed-loop temperature control system (Eurotherm 3216), with the sample surface temperature inside the chamber (measured *via* a K-type thermocouple) continuously used as the input. The chamber can accommodate up to two samples, which are illuminated using an unpolarized halogen white light source (AvaLight-HAL, Avantes) coupled through a bifurcated optical fiber (FCB-UV600-2, Avantes BV) equipped with collimating lenses. The transmitted light from each sample is collected and analyzed by a dual channel fiber-coupled fixed-grating spectrometer (AvaSpec-ULS2048CL-2-EVO, Avantes BV).

The samples were exposed to a constant gas flow rate of 200 mL/min of argon/H_2 mixtures throughout the entire measurement (at atmospheric pressure). The process started with an initialization stage that lasted almost 3 h, where the sensor temperature was set to 120 °C. During this time, the sensor was exposed ten times to 100 mbar H_2 in argon, in order to (i) desorb previously adsorbed H_2O from the sample surface and (ii) acquire a stable sensor baseline. The main stage of the experiments was performed at 30 °C and consisted of three identical sets of H_2 pulses. The partial pressures of H_2 in each set were 0.6, 1, 5, 10, 30, 60, and 100 mbar. Within the set, each H_2 pulse lasted for 300 s, followed by 600 s at 1 bar of argon gas pressure.

ACKNOWLEDGMENTS

We gratefully acknowledge funding from the Swedish Research Council (Nos. 2020-04935 and 2021-05072), the Swedish Foundation for Strategic Research (SIP21-0032), the Area of Advance Nano at Chalmers, and the Competence Centre TechForH2. The Competence Centre TechForH2 is hosted by Chalmers University of Technology and is financially supported by the Swedish Energy Agency (P2021-90268) and the member companies Volvo, Scania, Siemens Energy, GKN Aerospace, PowerCell, Oxeon, RISE, Stena Rederi AB, Johnson Matthey, and Insplorion. We are also grateful for computational resources provided by the National Academic Infrastructure for Supercomputing in Sweden at NSC, PDC, and C3SE partially funded by the Swedish Research Council through grant agreement No. 2022-06725, the Berzelius resource provided by the Knut and Alice Wallenberg Foundation at NSC, and the Dutch National supercomputer Snellius hosted by SURF. Part of this work was carried out at the Chalmers MC2 cleanroom facility.

SUPPORTING INFORMATION

Supporting Information Available: Details on the Lorentzian fitting of the DFs; additional figures on Bayesian optimization convergence, size correction, single-wavelength and full-spectrum sensor optimization, and time-series optical response data.

* erhart@chalmers.se

[1] C. Langhammer, I. Zorić, B. Kasemo, and B. M. Clemens, Hydrogen Storage in Pd Nanodisks Characterized with a Novel Nanoplasmonic Sensing Scheme, Nano Lett. **7**, 3122 (2007).

[2] C. Langhammer, E. M. Larsson, B. Kasemo, and I. Zorić, Indirect nanoplasmonic sensing: Ultrasensitive experimental platform for nanomaterials science and optical nanocalorimetry, Nano Lett. **10**, 3529 (2010).

[3] N. Liu, M. L. Tang, M. Hentschel, H. Giessen, and A. P. Alivisatos, Nanoantenna-enhanced gas sensing in a single tailored nanofocus, Nat. Mater. **10**, 631 (2011).

[4] C. Boelsma, L. J. Bannenberg, M. J. van Setten, N.-J. Steinke, A. A. van Well, and B. Dam, Hafnium — an optical hydrogen sensor spanning six orders in pressure, Nat. Commun. **8**, 15718 (2017).

[5] F. A. A. Nugroho, I. Darmadi, V. P. Zhdanov, and C. Langhammer, Universal scaling and design rules of hydrogen-induced optical properties in Pd and Pd-alloy nanoparticles, ACS Nano **12**, 9903 (2018).

[6] X. She, Y. Shen, J. Wang, and C. Jin, Pd films on soft substrates: A visual, high-contrast and low-cost optical hydrogen sensor, Light: Sci. Appl. **8**, 4 (2019).

[7] F. A. A. Nugroho, I. Darmadi, L. Cusinato, A. Susarrey-Arce, H. Schreuders, L. J. Bannenberg, A. B. da Silva Fanta, S. Kadkhodazadeh, J. B. Wagner, T. J. Antosiewicz, A. Hellman, V. P. Zhdanov, B. Dam, and C. Langhammer, Metal–polymer hybrid nanomaterials for plasmonic ultrafast hydrogen detection, Nat. Mater. **18**, 489 (2019).

[8] T. Beni, N. Yamasaku, T. Kurotsu, N. To, S. Okazaki, T. Arakawa, A. Balčytis, G. Seniutinas, S. Juodkazis, and Y. Nishijima, Metamaterial for Hydrogen Sensing, ACS Sens. **4**, 2389 (2019).

[9] F. Sterl, N. Strohfeldt, S. Both, E. Herkert, T. Weiss, and H. Giessen, Design Principles for Sensitivity Optimization in Plasmonic Hydrogen Sensors, ACS Sens. **5**, 917 (2020).

[10] L. Bannenberg, H. Schreuders, and B. Dam, Tantalum-Palladium: Hysteresis-Free Optical Hydrogen Sensor Over 7 Orders of Magnitude in Pressure with Sub-Second Response, Adv. Funct. Mater. **31**, 2010483 (2021).

[11] I. Darmadi, F. A. A. Nugroho, and C. Langhammer, High-Performance Nanostructured Palladium-Based Hydrogen Sensors—Current Limitations and Strategies for Their Mitigation, ACS Sens. **5**, 3306 (2020).

[12] T. B. Flanagan and W. A. Oates, The Palladium-Hydrogen System, Annu. Rev. Mater. Sci. **21**, 269 (1991).

[13] C. Wadell, F. A. A. Nugroho, E. Lidström, B. Iandolo, J. B. Wagner, and C. Langhammer, Hysteresis-free nanoplasmonic pd–au alloy hydrogen sensors, Nano Lett. **15**, 3563 (2015).

[14] I. Darmadi, F. A. A. Nugroho, S. Kadkhodazadeh, J. B. Wagner, and C. Langhammer, Rationally Designed PdAuCu Ternary Alloy Nanoparticles for Intrinsically Deactivation-Resistant Ultrafast Plasmonic Hydrogen Sensing, ACS Sens. **4**, 1424 (2019).

[15] V. Martvall, H. Klein Moberg, A. Theodoridis, D. Tomeček, P. Ekborg-Tanner, S. Nilsson, G. Volpe, P. Erhart, and C. Langhammer, Accelerating Plasmonic Hydrogen Sensors for Inert Gas Environments by Transformer-Based Deep Learning, ACS Sens. **10**, 376 (2025).

[16] S. De Marchi, S. Núñez-Sánchez, G. Bodelón, J. Pérez-Juste, and I. Pastoriza-Santos, Pd nanoparticles as a plasmonic material: Synthesis, optical properties and applications, Nanoscale **12**, 23424 (2020).

[17] P. Ekborg-Tanner, J. M. Rahm, V. Rosendal, M. Bancerek, T. P. Rossi, T. J. Antosiewicz, and P. Erhart, Computational design of alloy nanostructures for optical sensing of hydrogen, ACS Appl. Nano Mater. **5**, 10225 (2022).

[18] F. A. A. Nugroho, P. Bai, I. Darmadi, G. W. Castellanos, J. Fritzsche, C. Langhammer, J. Gómez Rivas, and A. Baldi, Inverse designed plasmonic metasurface with parts per billion optical hydrogen detection, Nat. Commun. **13**, 5737 (2022).

[19] A. Theodoridis, C. Andersson, S. Nilsson, J. Fritzsche, and C. Langhammer, A Catalytic-Plasmonic Pt Nanoparticle Sensor for Hydrogen Detection in High-Humidity Environments, ACS Sens. **10**, 8983 (2025).

[20] J. Hong, S. Lee, J. Seo, S. Pyo, J. Kim, and T. Lee, A Highly Sensitive Hydrogen Sensor with Gas Selectivity Using a PMMA Membrane-Coated Pd Nanoparticle/Single-Layer Graphene Hybrid, ACS Appl. Mater. Interfaces **7**, 3554 (2015).

[21] I. Darmadi, A. Stolaś, I. Östergren, B. Berke, F. A. A. Nugroho, M. Minelli, S. Lerch, I. Tanyeli, A. Lund, O. Andersson, V. P. Zhdanov, M. Liebi, K. Moth-Poulsen, C. Müller, and C. Langhammer, Bulk-Processed Pd Nanocube–Poly(methyl methacrylate) Nanocomposites as Plasmonic Plastics for Hydrogen Sensing, ACS Appl. Nano Mater. , acsanm.0c01907 (2020).

[22] S. Luo, D. Wang, and T. B. Flanagan, Thermodynamics of Hydrogen in fcc Pd–Au Alloys, J. Phys. Chem. B **114**, 6117 (2010).

[23] A. D. Humphrey and W. L. Barnes, Plasmonic surface lattice resonances on arrays of different lattice symmetry, Phys. Rev. B **90**, 075404 (2014).

[24] P. Chen, N. T. Tran, X. Wen, Q. Xiong, and B. Liedberg, Inflection Point of the Localized Surface Plasmon Resonance Peak: A General Method to Improve the Sensitivity, ACS Sens. **2**, 235 (2017).

[25] E. Herkert, F. Sterl, N. Strohfeldt, R. Walter, and H. Giessen, Low-Cost Hydrogen Sensor in the ppm Range with Purely Optical Readout, ACS Sens. **5**, 978 (2020).

[26] S. Celiksoy, W. Ye, K. Wandner, K. Kaefer, and C. Sönnichsen, Intensity-Based Single Particle Plasmon Sensing, Nano Lett. **21**, 2053 (2021).

[27] U.S. Department of Energy, Energy Efficiency and Renewable Energy (EERE), Fuel Cell Technologies Office, Multi-Year Research, Development, and Demonstration Plan, 2011–2020: Section 3.7 Hydrogen Safety, Codes and Standards (2015).

[28] J. M. Rahm, J. Löfgren, E. Fransson, and P. Erhart, A tale of two phase diagrams: Interplay of Ordering and Hydrogen Uptake in Pd–Au–H, Acta Mater. **211**, 116893 (2021).

[29] E. Brochu, V. M. Cora, and N. d. Freitas, A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning (2010), arXiv:1012.2599.

[30] F. A. A. Nugroho, B. Iandolo, J. B. Wagner, and C. Langhammer, Bottom-Up Nanofabrication of Supported Noble Metal Alloy Nanoparticle Arrays for Plasmonics, ACS Nano **10**, 2871 (2016).

[31] J. M. Rahm, C. Tiburski, T. P. Rossi, F. A. A. Nugroho, S. Nilsson, C. Langhammer, and P. Erhart, A Library of Late Transition Metal Alloy Dielectric Functions for Nanophotonic Applications, Adv. Funct. Mater. **30**, 2002122 (2020).

[32] J. J. Mortensen, A. H. Larsen, M. Kuisma, A. V. Ivanov, A. Taghizadeh, A. Peterson, A. Haldar, A. O. Dohn, C. Schäfer, E. Ö. Jónsson, and K. S. Thygesen, GPAW: An Open Python Package for Electronic Structure Calculations, J. Chem. Phys. **160**, 092503 (2024).

[33] GPy, GPy: A gaussian process framework in python, `http://github.com/SheffieldML/GPy` (since 2012).

[34] T. Schaul, Benchmarking separable natural evolution strategies on the noiseless and noisy black-box optimization testbeds, in *Proceedings of the 14th annual conference companion on Genetic and evolutionary computation* (ACM, Philadelphia Pennsylvania USA, 2012) pp. 205–212.

[35] Lumerical Inc., Ansys Lumerical FDTD (2022), version 2022 R1.

[36] F. A. A. Nugroho, D. Albinsson, T. J. Antosiewicz, and C. Langhammer, Plasmonic Metasurface for Spatially Resolved Optical Sensing in Three Dimensions, ACS Nano **14**, 2345 (2020).

Supporting Information

Bayesian Optimization for Practical H_2 Sensors: Inverse Design of Pd-based Plasmonic Metasurfaces

Pernilla Ekborg-Tanner[1], Athanasios Theodoridis[1], Joachim Fritzsche[1], Christoph Langhammer[1], Andrea Baldi[2], and Paul Erhart[*,1,3]

[1] Department of Physics and Astronomy, Chalmers University of Technology SE-412 96 Gothenburg, Sweden

[2] Department of Physics and Astronomy, Vrije Universiteit Amsterdam De Boelelaan 1100, 1081 HZ Amsterdam, Netherlands

[3] Wallenberg Initiative Materials Science for Sustainability, Chalmers University of Technology SE-412 96 Gothenburg, Sweden

[*] erhart@chalmers.se

List of Abbreviations

DF dielectric function.

FoM figure of merit.

FWHM full width at half maximum.

PMMA poly(methyl methacrylate).

Contents

Supplementary Notes **S2**
SN1. Lorentzian representation fit of dielectric functions S2

Supplementary Tables **S3**
S1. Samples selected for experimental verification S3

Supplementary Figures **S4**
S1. Bayesian optimization for 10 mbar H_2 and the 650 nm single-wavelength FoM. S4
S2. Post-anneal size correction of extinction spectra. S5
S3. Single-wavelength sensor optimization. S5
S4. Effect of Lorentzian fit on single-disk sensitivity. S6
S5. Full-spectrum sensor optimization. S6
S6. Time series for the optical response of the full-spectrum sensors. S7
S7. Extended FoM comparison for all samples including the full-spectrum sensors. S7
S8. Time series for the optical response of the 450 nm single-wavelength sensors. S8
S9. Time series for the optical response of the 550 nm single-wavelength sensors. S8
S10. Time series for the optical response of the 650 nm single-wavelength sensors. S9
S11. Fitted dielectric functions and extinction spectra. S9
S12. Pure Pd inconsistency. S10

Supplementary References **S11**

Supplementary Notes

Supplementary Note 1: Lorentzian representation fit of dielectric functions

In order to allow for smooth changes of the DF with composition, the DFs are described as a sum of Lorentzians,

$$\epsilon(\omega) = 1 + \sum_{j=1}^{N} \frac{x_j \omega_j^2}{\omega_j^2 - \omega^2 - i\omega\gamma_j}, \tag{1}$$

where each Lorentzian is characterized by its frequency ω_j, strength x_j and broadening γ_j. Each of these $3N$ parameters is expressed as polynomial functions of composition and fitted to the calculated DFs in the three steps.

1. The calculated DFs are individually fitted to a Lorentzian representation using the method described in Ref. 1, resulting in 5 Lorentzians per DF.

2. The individually fitted Lorentzians are sorted by frequency, and each of the three parameters is fitted for each Lorentzian to a second-order polynomial of the H and Au concentrations, resulting in a total of 15 fitted polynomials.

3. The polynomials from step 2 are used as initial guesses in a final NLLS fit (as implemented in SciPy (2)) of all coefficients of the 15 polynomials at once (90 coefficients in total). The residual in this fit is defined as the summed error of the real and imaginary parts of the DF at discrete energies with an added penalty for negative frequencies and broadening factors.

In Figure S11 examples of fitted vs. as-calculated DFs are shown as well as resulting extinction spectra. Note that the DF of pure Pd was intentionally skipped since it deviates from the trend of all other DFs (Figure S12). We believe this is due to some numerical issue in GPAW (3) that we were not able to resolve. The fitting procedure clearly results in smooth DFs with the same general shape but less fine structure compared to the as-calculated DFs. We have found that the resulting single-disk spectra reproduce experimentally observed trends (4) very well (Figure S4), in contrast to previous extensive attempts considering size distribution and artificial broadening. This indicates that enforcing smooth variations of the DF with composition and effectively removing most fine structure results in more realistic DFs.

Supplementary Tables

Table S1: Samples selected for experimental verification Overview of the target pressure, FoM and design parameters for all samples selected for experimental verification. All parameters are rounded from the predicted optimum values, with disk dimensions and array pitch rounded to the nearest 5 nm, the PMMA thickness to the nearest 25 nm, and the Au concentration to the nearest 5 %. The rounding precision reflects the estimated fabrication uncertainty and the width of the predicted FoM maximum for each parameter.

Sample	Pressure (mbar)	FoM	Disk radius (nm)	Disk height (nm)	Array pitch (nm)	PMMA (nm)	Au conc. (%)
B1	100	$\Delta\lambda$/FWHM	70	20	510	75	0
B2	10	$\Delta\lambda$/FWHM	65	70	300	300	15
B3	1	$\Delta\lambda$/FWHM	75	70	285	300	30
C1	100	Δext. @ 450 nm	50	25	200	175	0
C2	10	Δext. @ 450 nm	70	20	200	175	15
C3	1	Δext. @ 450 nm	70	20	200	0	30
D1	100	Δext. @ 550 nm	80	25	235	250	0
D2	10	Δext. @ 550 nm	75	20	230	225	15
D3	1	Δext. @ 550 nm	75	20	200	125	30
E1	100	Δext. @ 650 nm	85	20	315	75	0
E2	10	Δext. @ 650 nm	90	20	280	75	15
E3	1	Δext. @ 650 nm	90	20	260	300	30

Supplementary Figures

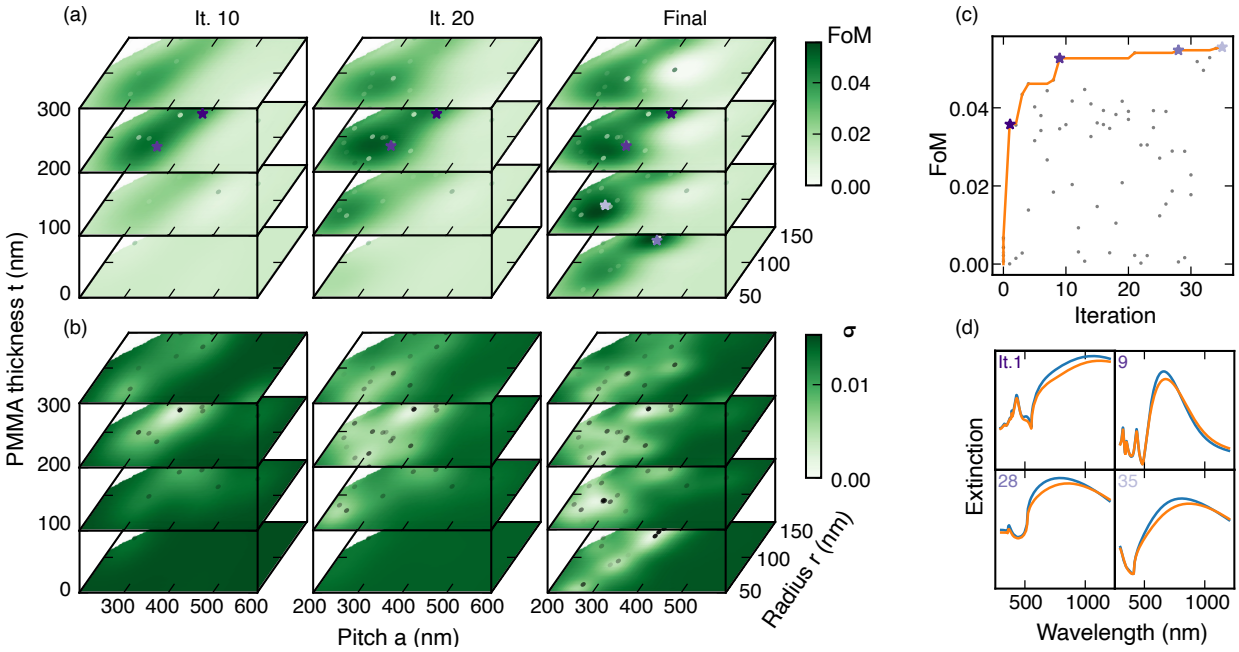

Figure S1: Bayesian optimization for 10 mbar H_2 and the 650 nm single-wavelength FoM. (a) Slices of the predicted FoM landscape and (b) the corresponding standard deviation as a function of array pitch, disk radius and PMMA thickness as it develops over the iterations (columns). For the selected slices, the disk height is 20 nm and the Au concentration is 15 %. The data points are included as dots with the inverse color scale and their opacity reflecting the closeness to the slice. (c) The obtained FoMs during the iterations and (d) the extinction spectra of the geometries with highest FoM at four selected iterations (indicated by stars in (a, c)).

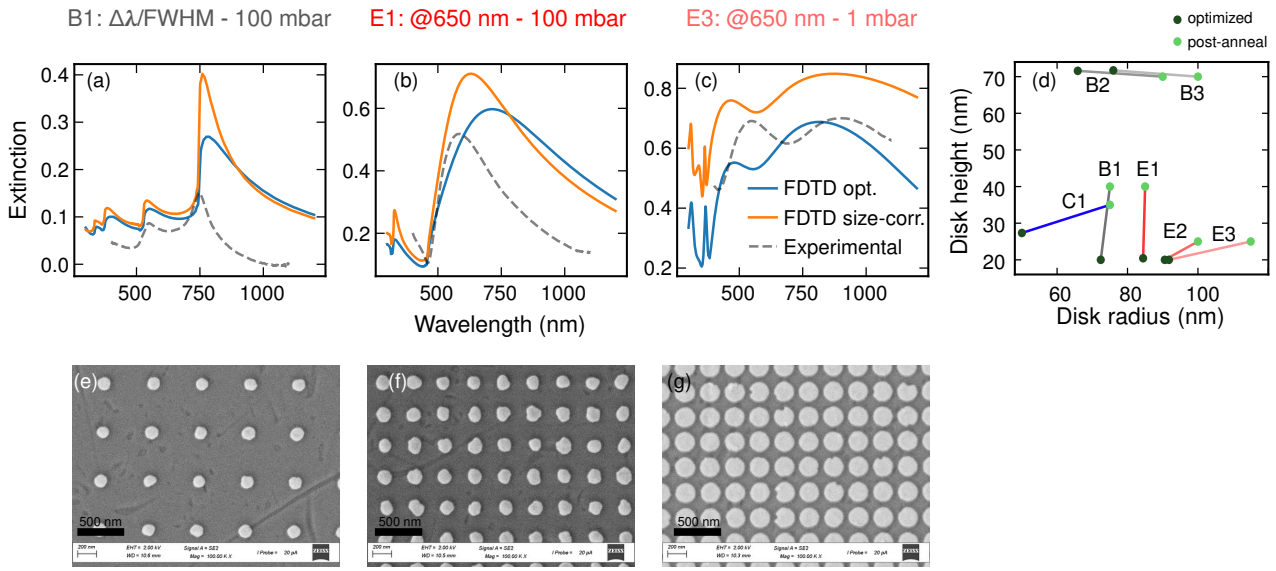

Figure S2: Post-anneal size correction of extinction spectra. The fabrication process is associated with a large uncertainty in the final size parameters of the nanodisks. (a-c) Optimized and size-corrected simulated extinction spectra for three example samples. (d) The change in the disk height and radius after the fabrication process (post-anneal) compared to the optimized parameters. (e-g) scanning electron microscope images of the fabricated samples (post-anneal). The first column corresponds to the sample optimized for peak shift over FWHM at 100 mbar (left) followed by samples optimized for extinction change at $\lambda = 650$ nm and 100 mbar (middle) or 1 mbar (right).

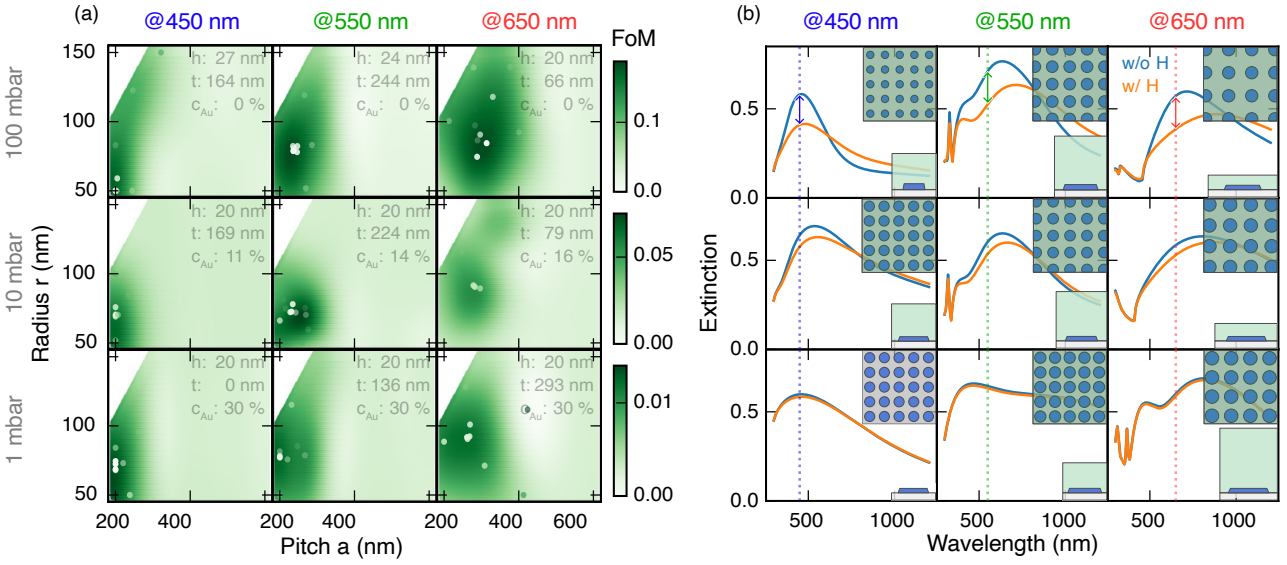

Figure S3: Single-wavelength sensor optimization. Optimization of the sensor geometries for target pressures 100 mbar (first row), 10 mbar (second row) and 1 mbar (third row). The FoM is based on the extinction change at a single wavelength, chosen to be 450, 550 and 650 nm (as indicated by the column titles). (a) Slices of the predicted FoM landscape surrounding the global maxima (specified by the annotated text). (b) Calculated extinction spectra in the pristine (blue) and hydrogenated states (orange) as well as schematic top and side views of the optimized samples.

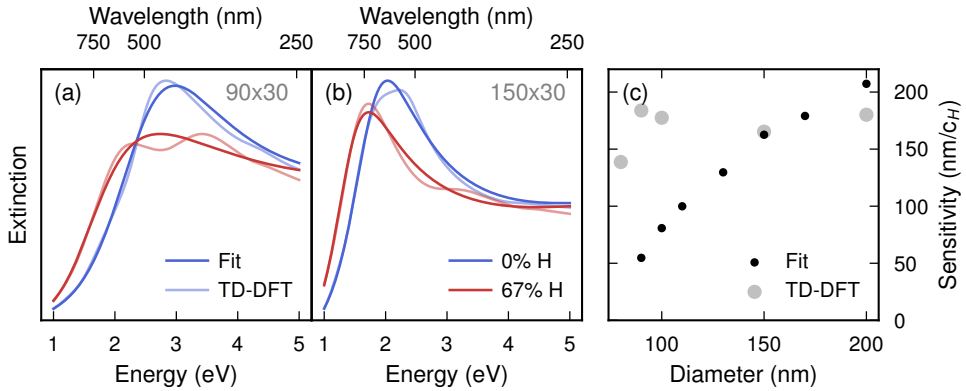

Figure S4: Effect of Lorentzian fit on single-disk sensitivity. The Lorentzian fit generally smooths features in the DFs and, in turn, in extinction spectra. Specifically, the double-peak phenomenon observed in our previous work (1) is smoothed out. This leads to a continuously increasing sensitivity with the disk diameter which is in better agreement with experimental observations (4).

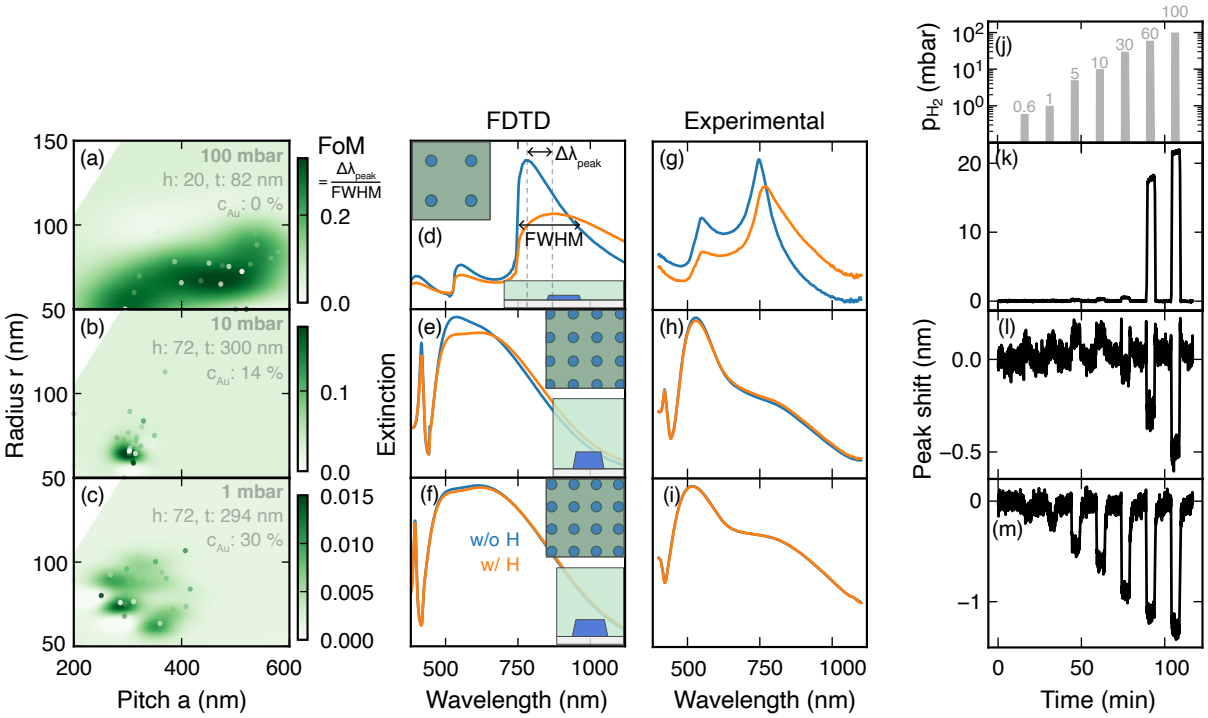

Figure S5: Full-spectrum sensor optimization. Optimization of the sensor geometries for target pressures (a) 100 mbar, (b) 10 mbar and (c) 1 mbar with peak shift over FWHM as FoM. (a-c) Slices of the FoM landscape surrounding the global maxima. (d-f) The simulated extinction spectra with H (blue) and without H (orange) as well as schematic top and side views of the optimized samples. (g-i) The measured extinction spectra. (k-m) The time evolution of the peak position as the samples are exposed to a sequence of hydrogen pulses ranging from 0.6 mbar to 100 mbar (j).

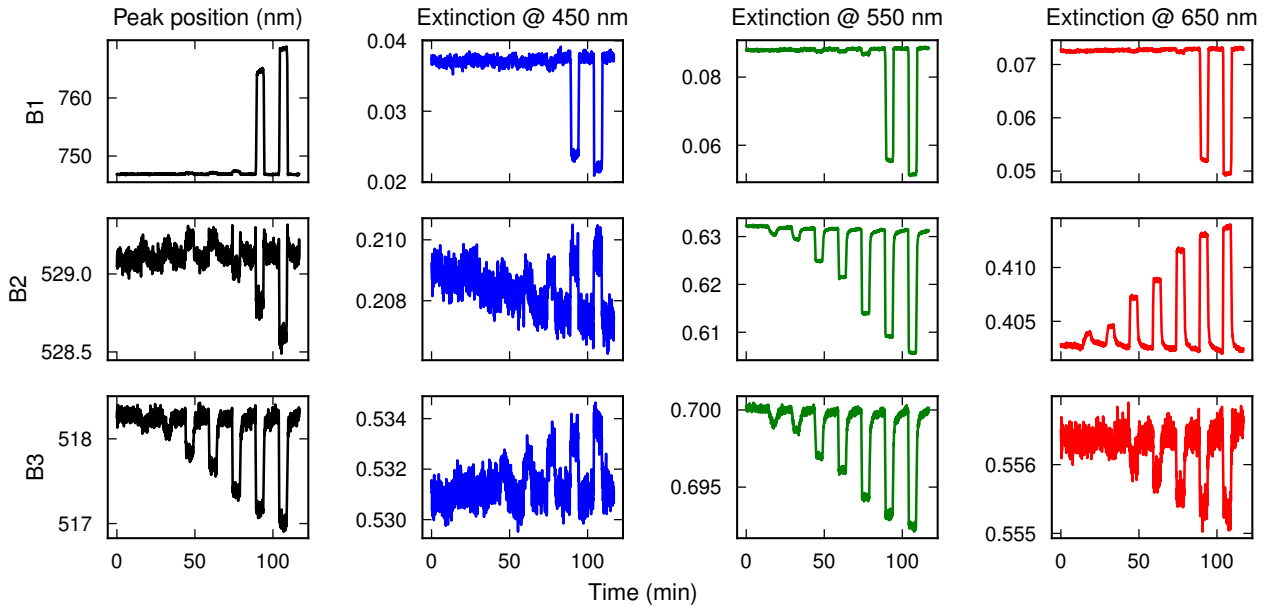

Figure S6: Time series for the optical response of the full-spectrum sensors. Time series for the optical response of the full-spectrum sensors, with decreasing target pressure from the top down, at all four FoMs.

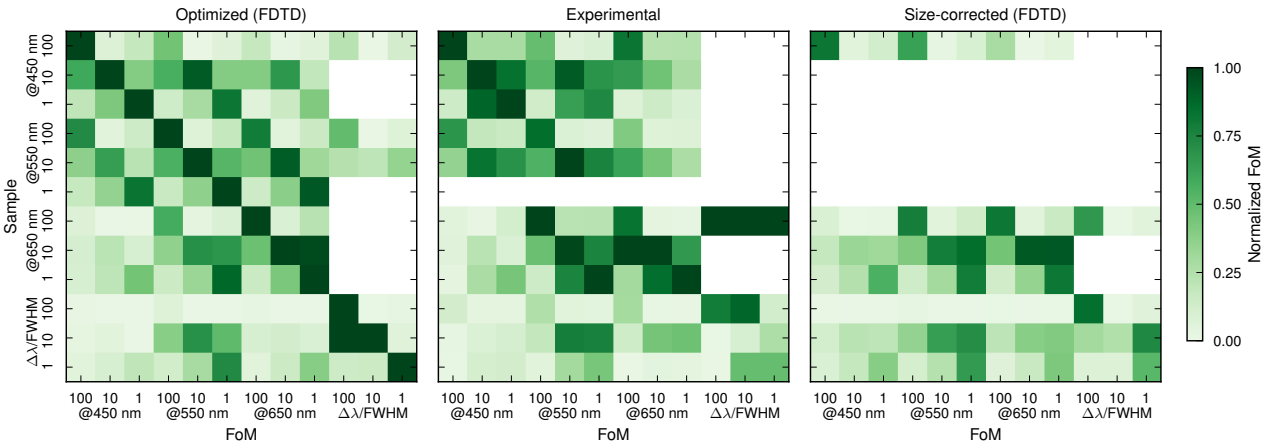

Figure S7: Extended FoM comparison for all samples including the full-spectrum sensors. All considered FoMs for all samples based on the simulations of the optimized samples (left), experimental measurements (middle) and simulations of the (post-anneal) size-corrected samples (right). Note that only a subset of the samples has been characterized with scanning electron microscopy, and that the size-corrected comparison uses the same normalization as for the optimized samples.

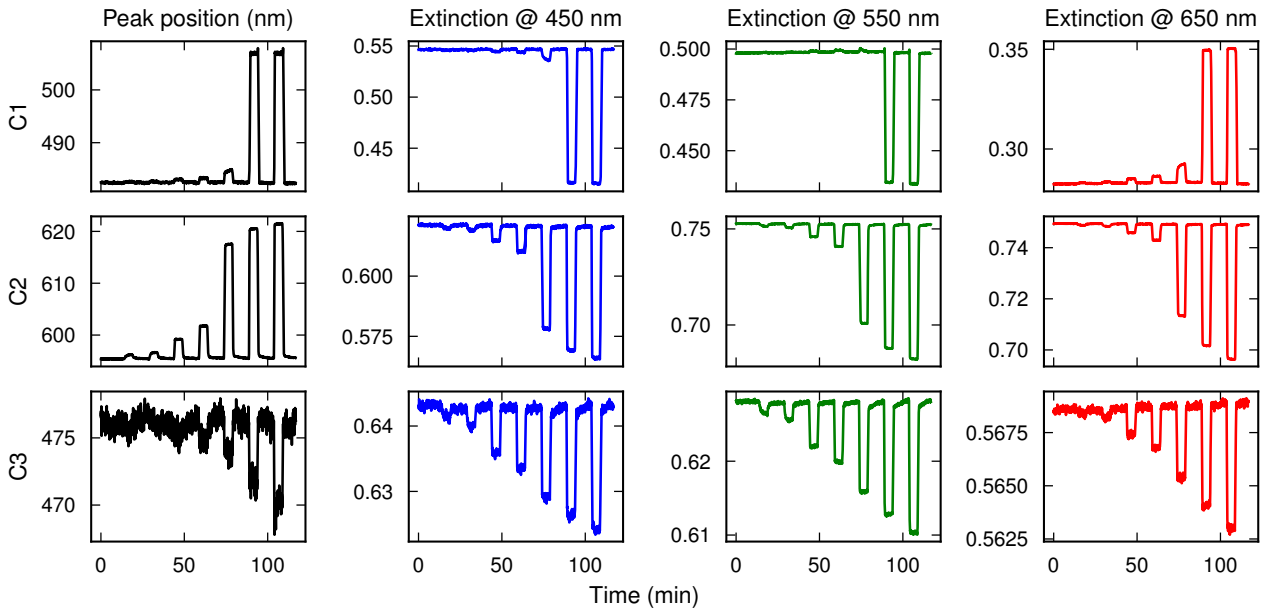

Figure S8: Time series for the optical response of the 450 nm single-wavelength sensors. Time series for the optical response of the 450 nm single-wavelength sensors, with decreasing target pressure from the top down, at all four FoMs.

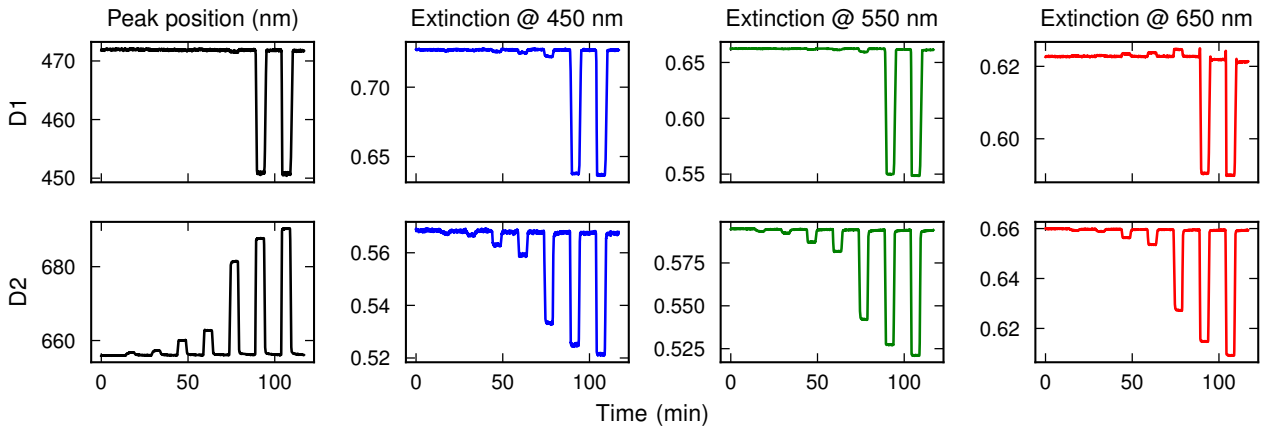

Figure S9: Time series for the optical response of the 550 nm single-wavelength sensors. Time series for the optical response of the 550 nm single-wavelength sensors, with decreasing target pressure from the top down, at all four FoMs.

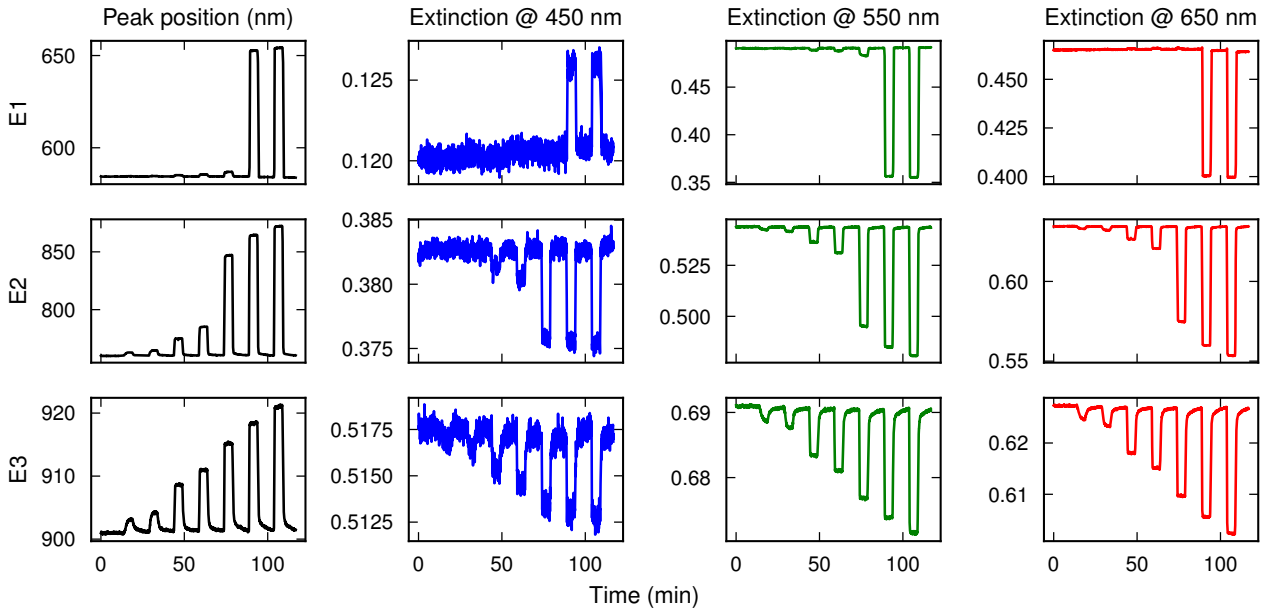

Figure S10: Time series for the optical response of the 650 nm single-wavelength sensors. Time series for the optical response of the 650 nm single-wavelength sensors, with decreasing target pressure from the top down, at all four FoMs.

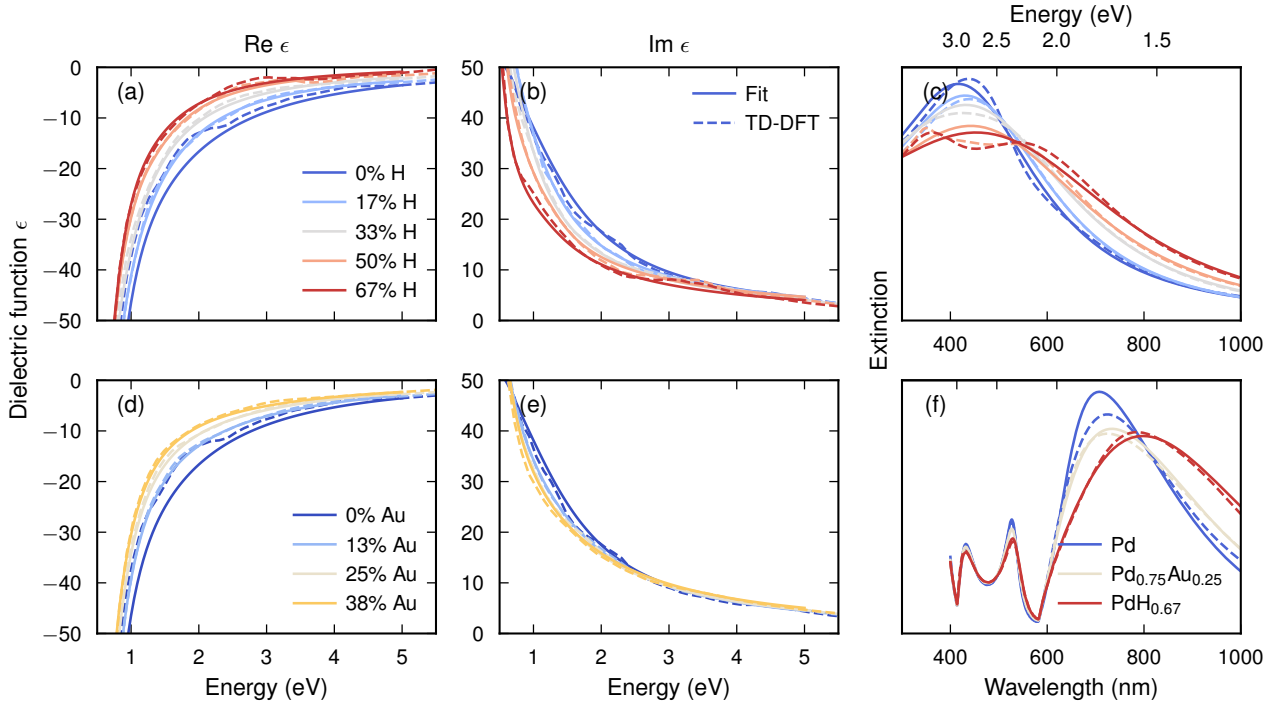

Figure S11: Fitted dielectric functions and extinction spectra. Real (a, d) and imaginary (b, e) parts of the dielectric function as-calculated with TD-DFT and after fitting for PdH (a, b) and PdAu (d, e). (c) Extinction spectra for single disks with diameter 90 nm and height 30 nm. (f) Extinction spectra for ordered nanodisk arrays with diameter 160 nm, height 40 nm, pitch 400 nm and PMMA thickness 300 nm.

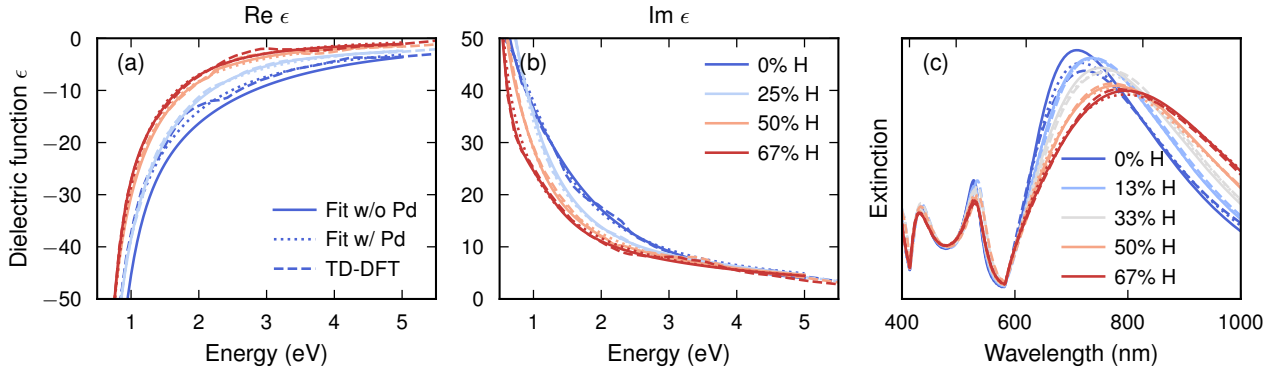

Figure S12: Pure Pd inconsistency. The dielectric function of pure Pd deviates from the trend of all other compositions. This is visible as a shoulder in the real part at $\sim 2\,\mathrm{eV}$ (a) and a lower amplitude than expected in the plasmon peak (c). We have not been able to find any support for that behavior in the literature and believe it to be an artifact of the density functional theory calculations. For this reason, we have chosen to exclude pure Pd when fitting the final Lorentzian representation.

Supplementary References

[1] P. Ekborg-Tanner, J. M. Rahm, V. Rosendal, M. Bancerek, T. P. Rossi, T. J. Antosiewicz, and P. Erhart, *Computational Design of Alloy Nanostructures for Optical Sensing of Hydrogen,* ACS Appl. Nano Mater. **5**, 10225 (2022). doi:10.1021/acsanm.2c01189.

[2] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python,* Nat. Methods **17**, 261 (2020). doi:10.1038/s41592-019-0686-2.

[3] J. J. Mortensen, A. H. Larsen, M. Kuisma, A. V. Ivanov, A. Taghizadeh, A. Peterson, A. Haldar, A. O. Dohn, C. Schäfer, E. Ö. Jónsson, and K. S. Thygesen, *GPAW: An Open Python Package for Electronic Structure Calculations,* J. Chem. Phys. **160**, 092503 (2024). doi:10.1063/5.0182685.

[4] F. A. A. Nugroho, I. Darmadi, V. P. Zhdanov, and C. Langhammer, *Universal Scaling and Design Rules of Hydrogen-Induced Optical Properties in Pd and Pd-Alloy Nanoparticles,* ACS Nano **12**, 9903 (2018). doi:10.1021/acsnano.8b02835.