(19) **United States**

(12) **Patent Application Publication** (10) Pub. No.: **US 2022/0407001 A1**

Takeuchi et al. (43) **Pub. Date:** **Dec. 22, 2022**

(54) **NOVEL NANOCOMPOSITE PHASE-CHANGE MEMORY MATERIALS AND DESIGN AND SELECTION OF THE SAME**

(71) Applicants: **University of Maryland, College Park**, College Park, MD (US); **University of Washington**, Seattle, WA (US); **Government of the United States of America as represented by the Secretary of Commerce**, Gaithersburg, MD (US)
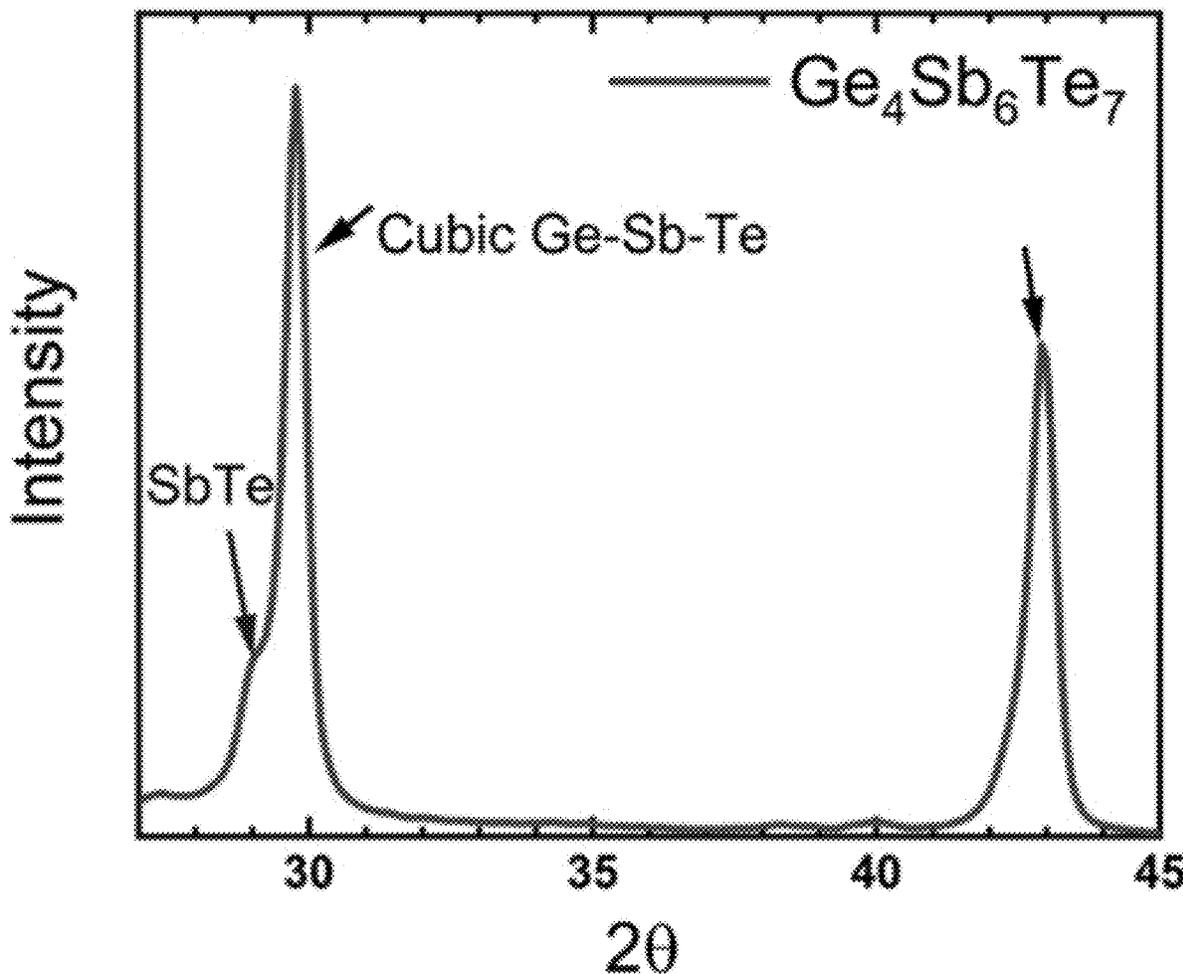
(72) Inventors: **Ichiro Takeuchi**, Laurel, MD (US); **Heshan Yu**, Greenbelt, MD (US); **Changming Wu**, Seattle, WA (US); **Mo Li**, Seattle, WA (US); **Aaron Gilad Kusne**, Rockville, MD (US)

(21) Appl. No.: **17/651,946**

(22) Filed: **Feb. 22, 2022**

Related U.S. Application Data

Publication Classification

(57) **ABSTRACT**

Provided herein are novel materials, such as novel phase-change memory materials providing superior characteristics, and methods of discovering/selecting such novel materials via machine learning, such as Bayesian active learning. An exemplary material provided by the inventive concept is the nanocomposite phase-change memory material $Ge_4Sb_6Te_7$, selected using closed-loop autonomous materials exploration and optimization (CAMEO).

FIG. 1

FIG. 2

FIG. 3

FIG. 4

FIG. 5

FIG. 6

FIG. 7

FIG. 7 (cont'd.)
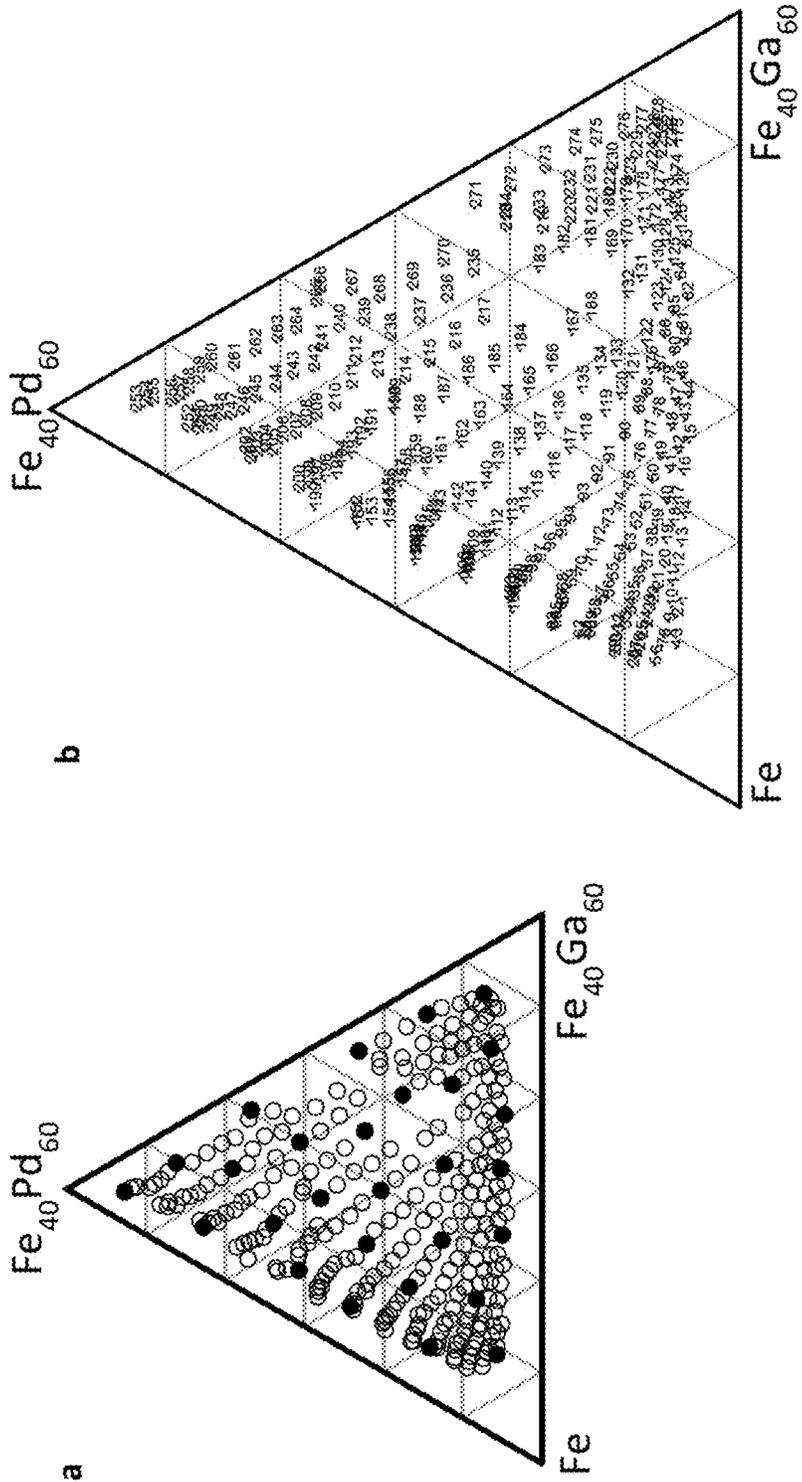
FIG. 8

FIG. 9

FIG. 10

FIG. 11

FIG. 12

FIG. 13

% of Simulations Runs Reaching within 1 % of Optimal

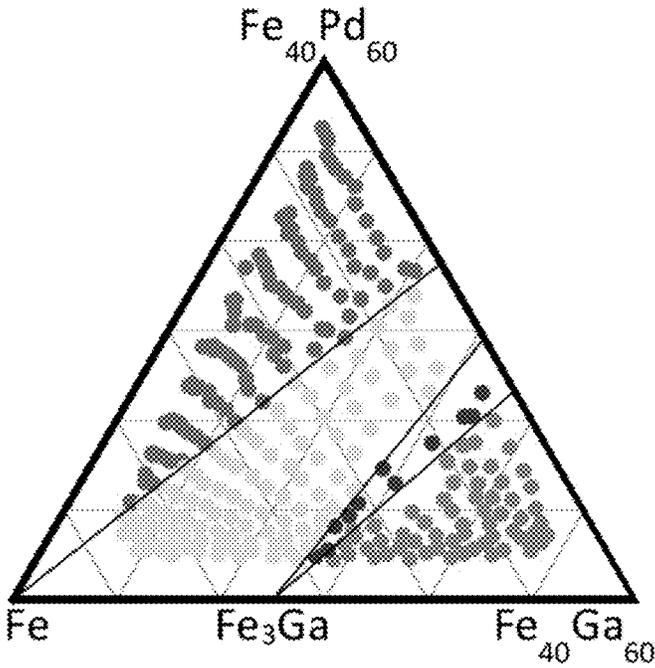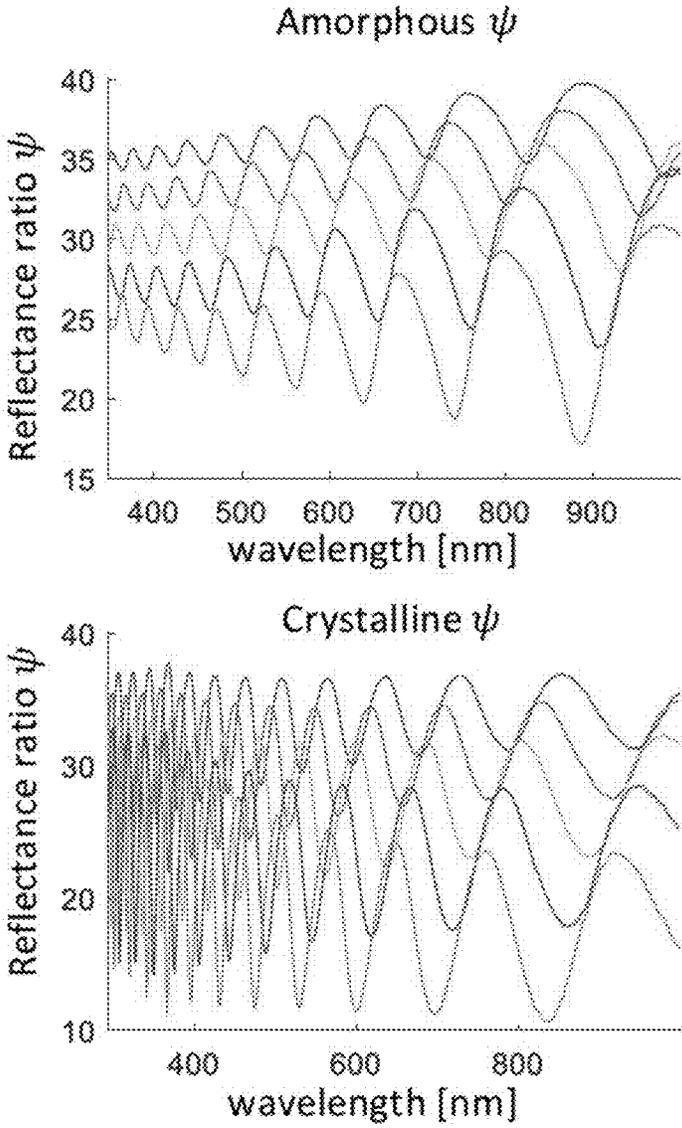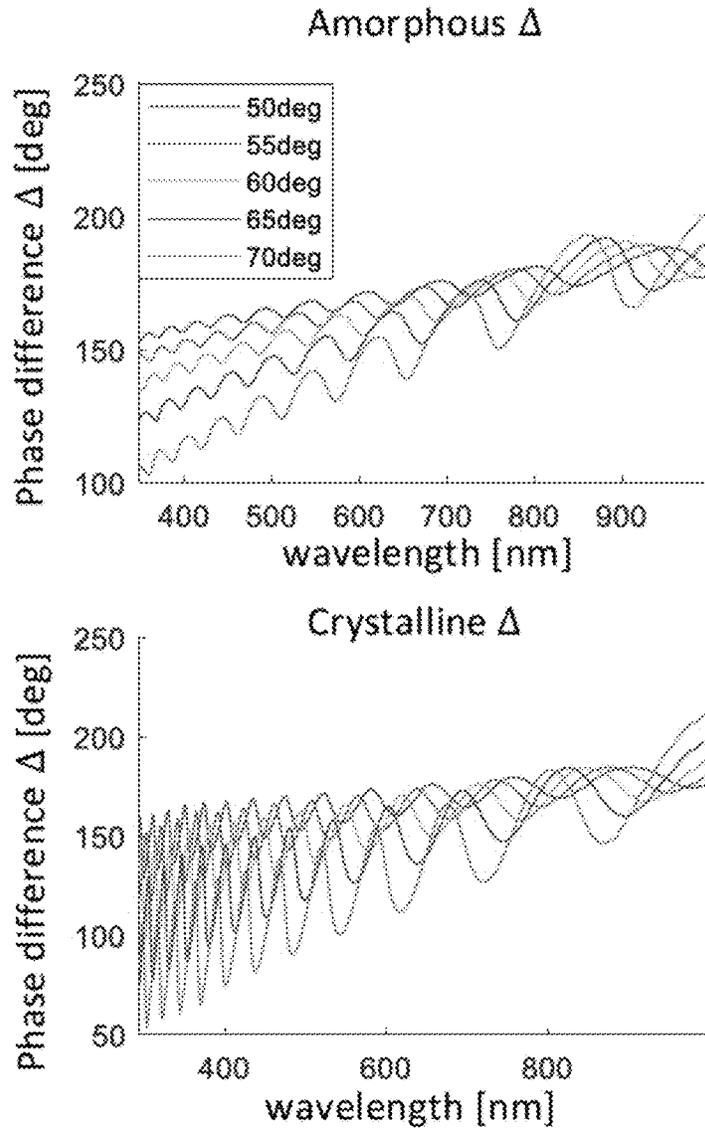FIG. 14

FIG. 15

FIG. 16

A

Refractive Index

Amorphous
Crystalline

8

7

6

5

4

1000 1200 1400 1600
λ (*nm*)

Extinction coefficient
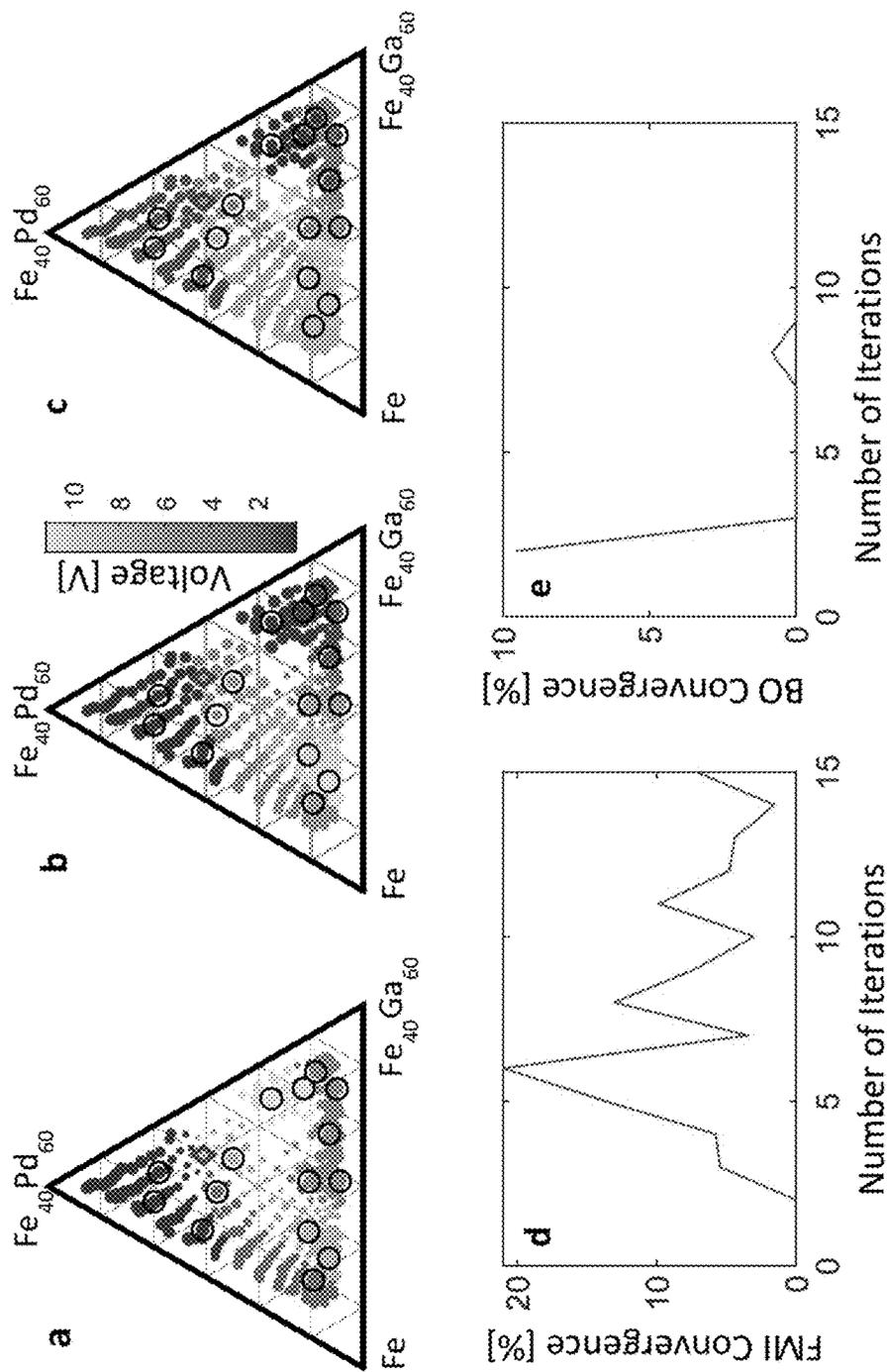
4

3

2

1

0

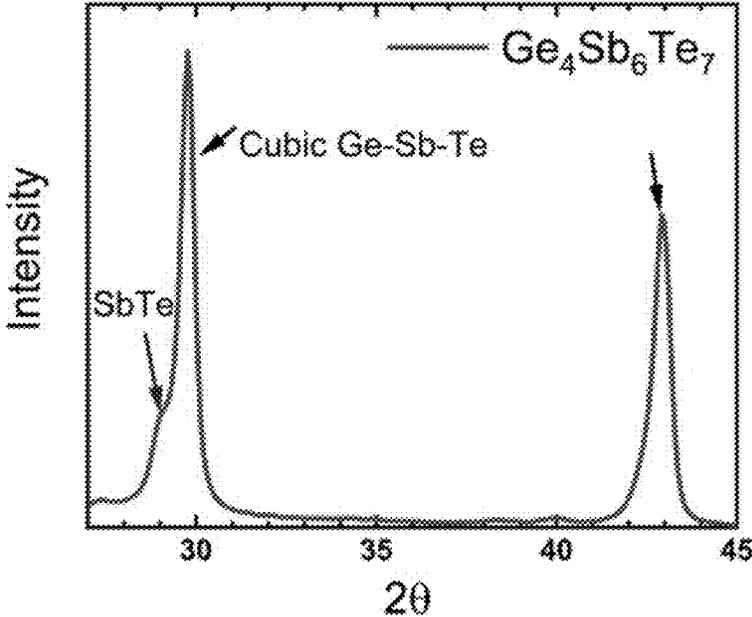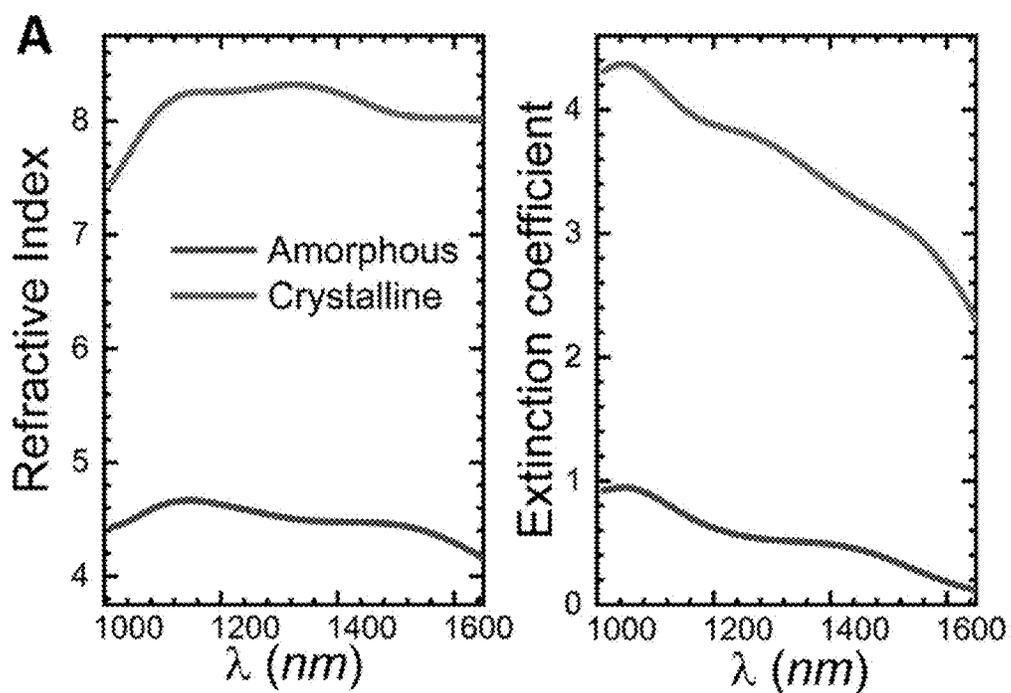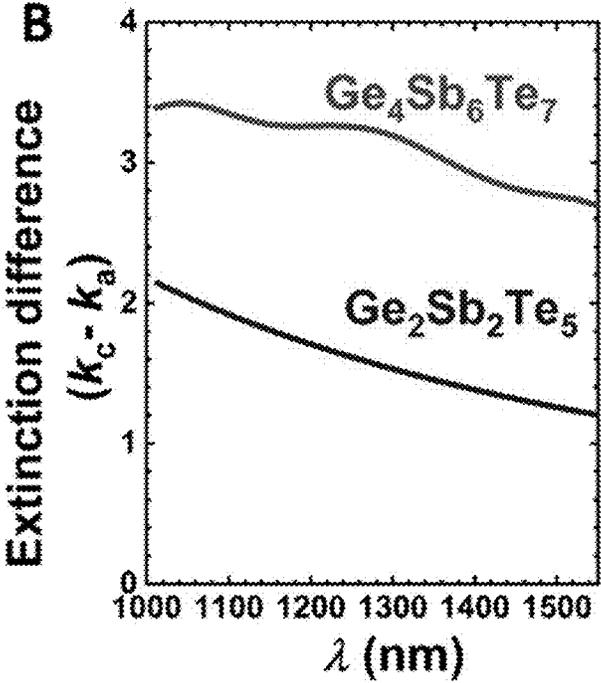1000 1200 1400 1600
λ (*nm*)
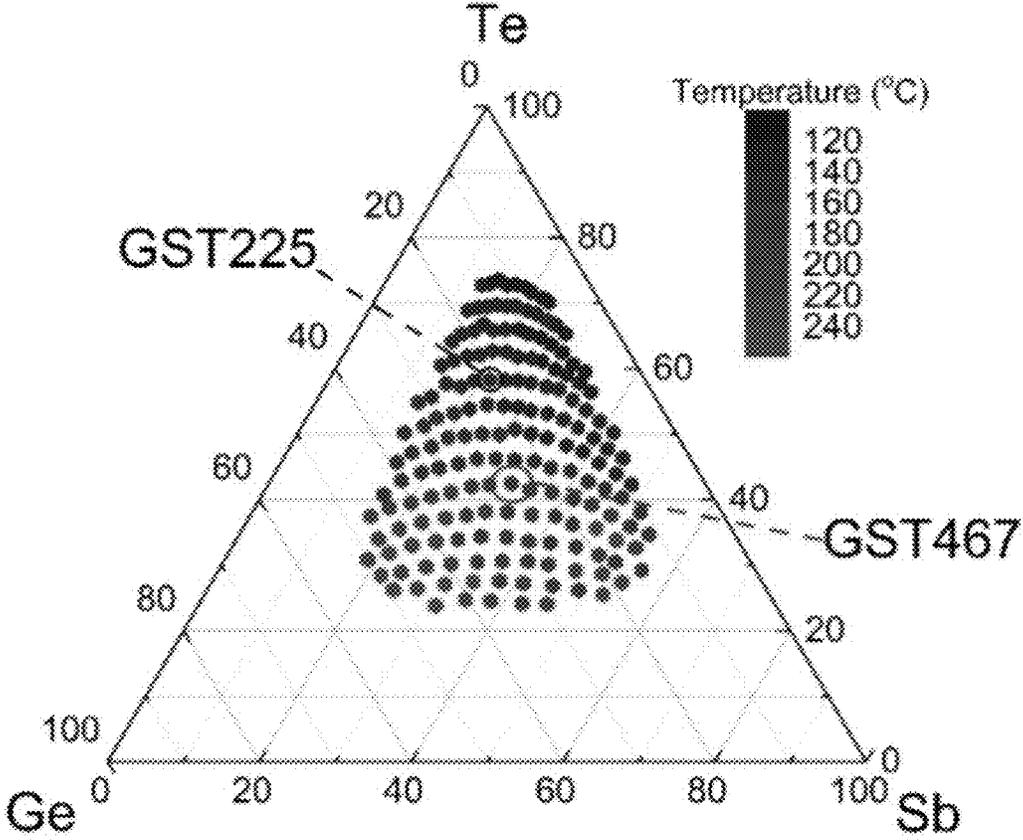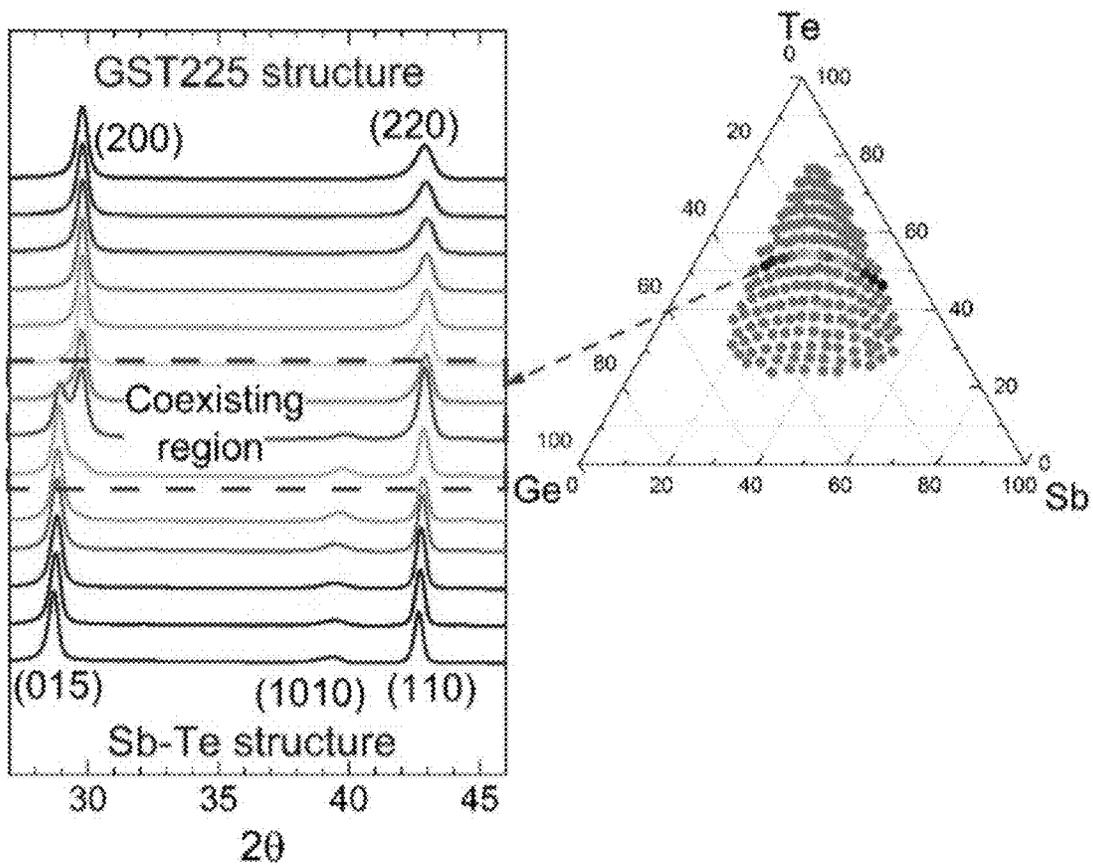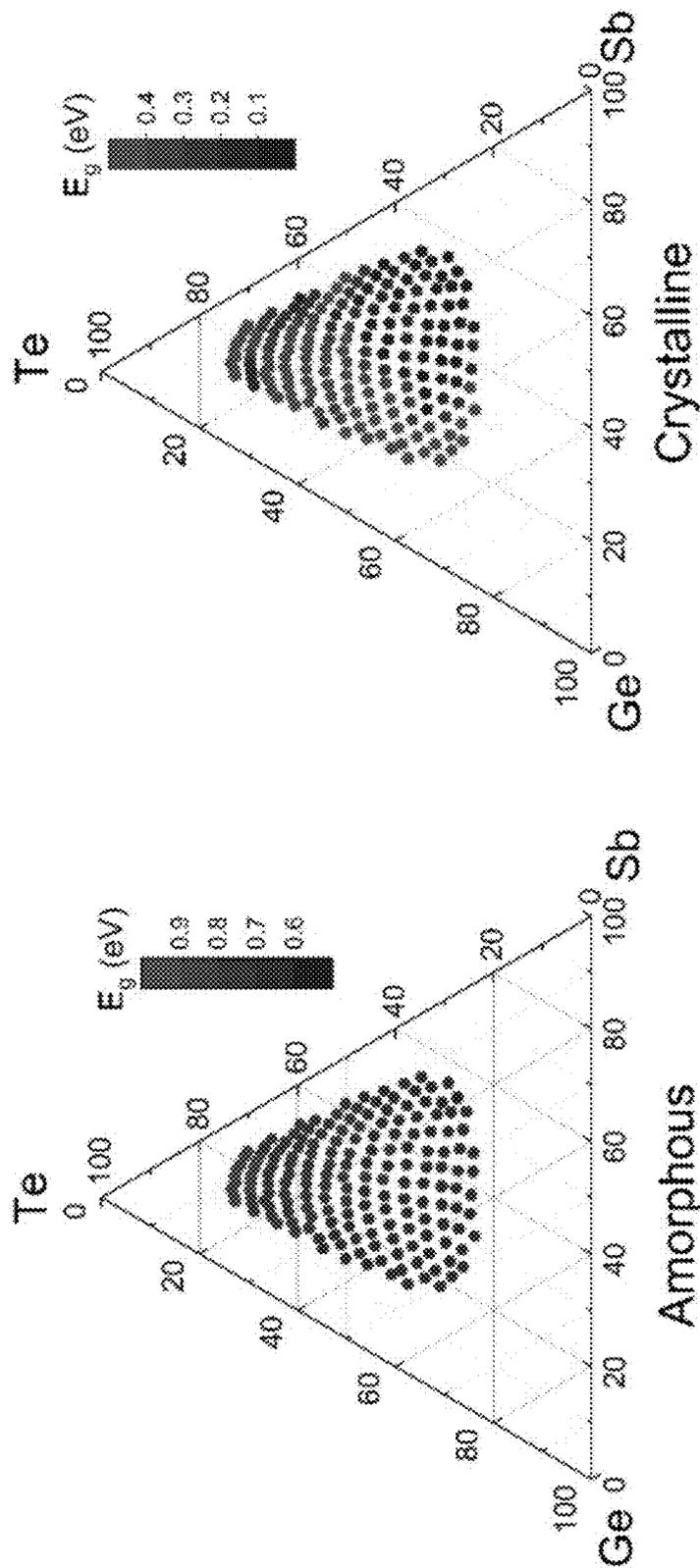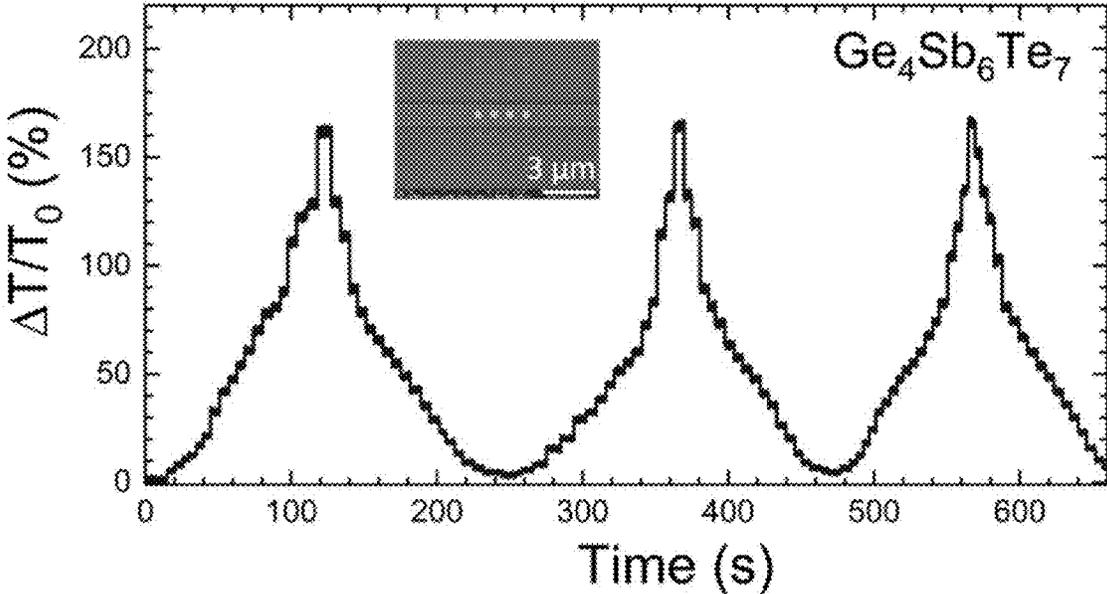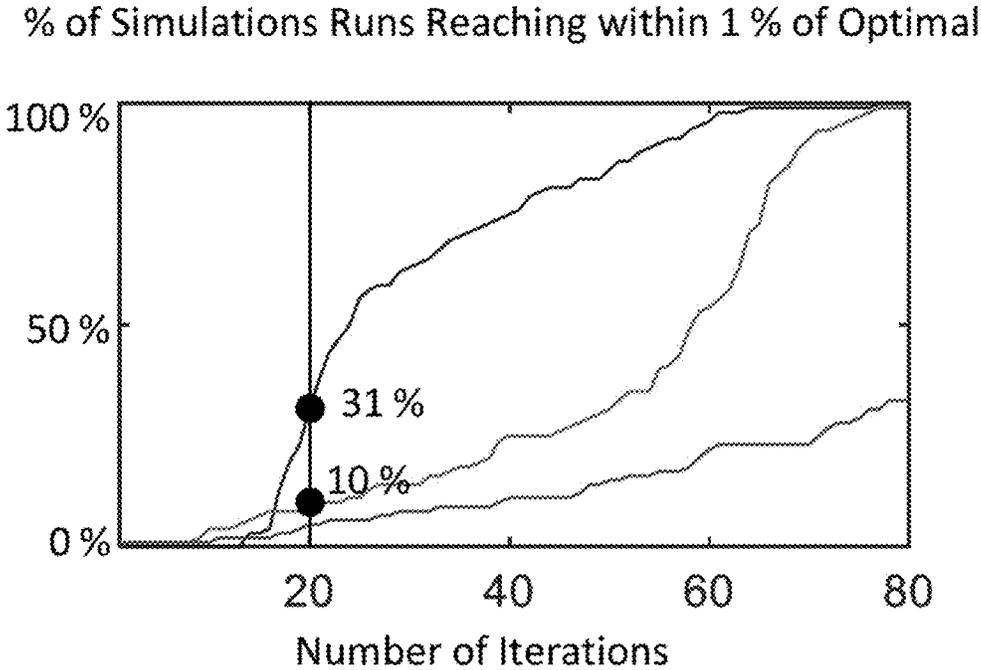
FIG. 17

FIG. 17 (cont'd.)

NOVEL NANOCOMPOSITE PHASE-CHANGE MEMORY MATERIALS AND DESIGN AND SELECTION OF THE SAME

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to U.S. Provisional Application Ser. No. 63/151,263, filed Feb. 19, 2021, the disclosure of which hereby is incorporated herein by reference in its entirety.

STATEMENT OF GOVERNMENT SUPPORT

FIELD

[0003] The subject matter of the present disclosure generally relates to improved materials, more particularly improved phase-change memory materials for use in the field of electronics and semiconductors, and the development and design of the same.

BACKGROUND

[0004] Technologies drive the perpetual search for novel and improved functional materials, necessitating the exploration of increasingly complex multi-component materials [1]. With each new component or material(s) parameter, the space of candidate experiments grows exponentially. For example, if investigating the impact of a new parameter (e.g., introducing doping) involves approximately ten experiments over the parameter range, N parameters will require on the order of 10^{N+} possible experiments. High-throughput synthesis and characterization techniques offer a partial solution: with each new parameter, the number of candidate experiments rapidly escapes the feasibility of exhaustive exploration. The search is further confounded by the diversity and complexity of materials composition-structure-property (CSP) relationships, including materials-processing parameters and atomic disorder [2]. Coupled with the sparsity of optimal materials, these challenges threaten to impede innovation and industrial advancement.

[0005] Structural phase maps, which describe the dependence of materials structure on composition, serve as blueprints in the design of functional and structural materials, as most materials properties are tied to crystal-structure prototypes. For example, property extrema tend to occur within specific phase regions (e.g., magnetism and superconductivity) or along phase boundaries (e.g., caloric-cooling materials and morphotropic phase-boundary piezoelectrics). Structural phase maps, and more specifically equilibrium phase diagrams, were traditionally generated over years with point-by-point Edisonian approaches guided by expert knowledge and intuition and involving iterative materials synthesis, diffraction-based structure characterization, and crystallographic refinement.

[0006] Machine learning (ML) is transforming materials research before our eyes [3], and yet direct coupling of ML with experiments remains a formidable challenge. Closed-loop autonomous system for materials exploration and optimization (CAMEO) offers a new materials research paradigm to truly harness the accelerating potential of ML, setting the stage for the 21st-century paradigm of materials research—the autonomous materials research lab run under the supervision of a robot scientist or artificial scientist [4]. Furthermore, CAMEO embraces aspects of human-in-the-loop autonomous systems [5, 6], where the human provides their expertise while ML presides over decision making steps. Live visualization of data analysis and decision making (including uncertainty quantification) provides interpretability of the autonomous process for the human expert in the human-machine research team (see FIG. 15). CAMEO also exploits (as of yet) non-automated capabilities of the human expert in the closed loop, thus elevating the capabilities of both human and machine.

[0007] Active learning [7]—the ML field dedicated to optimal experiment design (i.e., adaptive design), is key to this new paradigm. Active learning provides a systematic approach to identify the best experiments to perform next to achieve user-defined objectives. Scientific application can be traced as far back as the 18th century to Laplace's guided discoveries of celestial mechanics [8]. Bayesian optimization (BO) active learning techniques have been used more recently to guide experimentalists in the lab to optimize unknown functions [9-14]. BO methods balance the use of experiments to explore the unknown function with experiments that exploit prior knowledge to identify extrema. However, these past studies only advised researchers on the next experiment to perform, leaving experiment planning, execution, and analysis to the researcher. Recent advances in robotics have shifted the burden of materials synthesis from human experts to automated systems, accelerating materials discovery [15,16]. Concurrently, active learning has been demonstrated to accelerate property optimization by guiding simulations of known phases [17]. More recently, autonomous systems and machine learning driven research have been demonstrated for optimizing process and system operation [18-20] sample characterization [21], and tuning chemical reactions of known polymers and organic molecules for technological applications [22-24], using off-the-shelf optimization schemes. Taking another step and placing active learning in real-time control of solid-state materials exploration labs promises to accelerate materials discovery while also rapidly and efficiently illuminating complex materials-property relationships. Such potential innovation has been discussed in recent perspectives [25,26], with a primary focus on autonomous chemistry [27-29].

SUMMARY

[0008] Herein, CAMEO (see FIG. 1) is demonstrated in both real-time control of X-ray diffraction measurement experiments over composition spreads at the synchrotron beamline and in the lab. The algorithm accelerates phase mapping and discovery of novel materials, such as novel solid-state materials, with a 10-fold reduction in required experiments, each iteration taking tens of seconds to tens of minutes depending on the experimental task.

[0009] According to an aspect, provided is a phase-change memory material including $Ge_4Sb_6Te_7$.

[0010] According to another aspect, provided is a phase-change memory material including a material of formula (I):

$$TM_xSb_yTe_z \qquad (I)$$

wherein TM is a transition metal, wherein x, y, and z are each independently an integer between 1-10, and wherein the phase-change memory material has an extinction difference between crystalline and amorphous phases $(k_c–k_a)$ greater than about 2 in a wavelength range of about 1000 nm to about 1500 nm.

[0011] According to another aspect, provided is a method of forming a photonic switching device including forming a film on a substrate, the film including a phase-change memory material selected to have an extinction difference between crystalline and amorphous phases $(k_c–k_a)$ greater than about 2 in a wavelength range of about 1000 nm to about 1500 nm.

[0012] According to another aspect, provided is a method of forming a memory device including forming a film on a substrate, the film including a phase-change memory material selected to have an extinction difference between crystalline and amorphous phases $(k_c–k_a)$ greater than about 2 in a wavelength range of about 1000 nm to about 1500 nm.

[0013] According to another aspect, provided is a method of selecting a phase-change memory material including: evaluating optical bandgap difference of phase-change memory materials; and selecting phase-change memory material or materials having larger or largest optical bandgap differences, wherein the evaluation is conducted via Bayesian active learning.

[0014] According to another aspect, provided is a method including: evaluating optical bandgap of a phase-change memory material via Bayesian active learning; and selecting a phase-change memory material or materials having larger or largest optical bandgap differences.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] FIG. 1: Closed-loop autonomous materials exploration and optimization (CAMEO). The autonomous cycle begins with loading data from databases including composition data for the materials on the composition spread and computed materials data from the AFLOW.org[41] density functional theory database. The collected data is then used to begin analysis of the data using physics-informed Bayesian machine learning. This process extends knowledge of structure and functional property from materials with data to those without, predicting their estimated structure and functional property, along with prediction uncertainty. Physics-informed active learning is then used to identify the most informative next material to study to achieve user-defined objectives. For this work, active learning can select the next sample to characterize through autonomous control of the high-throughput X-ray diffraction system at a synchrotron beamline or it can (optionally) request specific input from the human-in-the-loop. Future implementations will include autonomous materials synthesis and simulation. The data collected from measurements and from human input are added to the database and used for the next autonomous loop. For more information, see EXAMPLE 1, Methods.

[0016] FIG. 2: Comparison of materials optimization schemes. Panel a) Simple optimization seeks to identify the property optimum with a mixture of exploration and exploitation without knowledge of the composition-structure-property (CSP) relationship. These methods are more likely to get caught in local optima. Panel b) The phase-map-informed optimization scheme exploits CSP relationship by recognizing that the property is dependent on phase, thus including phase mapping in the search for the optimum. (i) Phase-mapping steps and ii) materials optimization step that exploits knowledge of the phase boundaries. This allows a search for phase region dependent optima. c The Ge—Sb—Te CAMEO workflow began with incorporating raw ellipsometry data as a phase-mapping prior. On each iteration, CAMEO selects a material to measure for X-ray diffraction and concurrently requests an expert to calculate ΔE_g for that material. Each cycle takes 20-25 min.

[0017] FIG. 3 Discovery of $Ge_4Sb_6Te_7$ (GST467). Panel a) Optimization of the Ge—Sb—Te system, with the objective of identifying the material with the largest ΔE_g, bandgap difference between amorphous and crystalline states. Performance is shown for: the CAMEO live run (black) with GST467 discovered on iteration 19 (black star); for the mean and 95% confidence interval of 100 CAMEO runs computed post data collection (blue), 100 runs of Gaussian process-upper confidence bounds (GP-UBC, magenta), and random sampling (red). For the 100 runs, CAMEO gets within 1% of the optimal in the first 20 runs 31% of the time compared to GP-UBC's 10%. Panel b) Left: structural phase map for the crystalline Ge—Sb—Te composition spread. Complete phase map constructed after all the diffraction measurements (beyond the live CAMEO run described in a)) is shown. Structural phase regions are color-coded as single-phase FCC—Ge—Sb—Te (GST) structure region (red, GST), single-phase Sb—Te region (blue), and region where GST and Sb—Te phases co-exist (orange). The materials measured during the live CAMEO run are indicated in black. Right: Complete mapping of ΔE_g following analysis of entire ellipsometry data (beyond the live CAMEO run). The discovered GST467 and GST225 are indicated in both maps.

[0018] FIG. 4 Nanostructure and device performance of $(Ge_4Sb_6Te_7)$ GST467. Newly identified phase-change memory material GST467 shows large optical contrast ideal for photonic-switching device applications such as neuromorphic computing. Panel a) High resolution transmission electron microscopy image reveals formation of coherent nanocomposite of GST structure matrix and SbTe. The dotted lines denote the atomically sharp interface. The FFT (inset) of this region indicates structural similarity of the adjacent phases; panel b) endurance of the GST467: it is stable over 30,000 cycles indicating the robustness of the nanocomposite structure defined by local composition variation. The dotted lines indicate the range of each state in relative optical transmission T/T_o at 1500 nm. Laser pulses were 50 ns with 183 pJ for quenching and 500 ns with 3.3 nJ for crystallization. The fluctuations in T/T_o are due to the thermal fluctuation of the device measurement set-up; panel c) comparison of the optical contrast here indicated by difference in the extinction coefficient k between crystalline and amorphous phases $(k_c–k_a)$ for the wavelength range of 1000-1500 nm for various compositions within Ge—Sb—Te system. GST225 and GST467 data are from this work. The GeTe data are from ref. 55. GST467 shows higher extinction difference over other known compositions; one-to-one comparison of panel d) GST225 (left) and panel e) GST467 (right) for multi-level switching in optical transmission at 1500 nm (T/T_o) using 500 ns, 6 pJ pulses: GST467 having larger optical contrast results in substan-

tially more states than GST225. Device fabrication and characterization details are described in EXAMPLE 1, Methods.

[0019] FIG. 5. Benchmarking performance. Panel a) Main figure: Phase mapping performance demonstrating that CAMEO provides a significant advantage over the three alternatives: random sampling, sequential sampling, and measuring 10% of the samples well distributed over the composition space. Subset figure: Material optimization performance. The benchmark materials optimization challenge is highly simple with a very prominent, broad peak—a challenge that Bayesian optimization schemes excel at. Nevertheless, CAMEO provides improved results over the next best alternative, GP-UCB. Of note is CAMEO's initial lag in performance due to its initial goal of maximizing phase mapping performance. Once phase mapping performance converges, it then switches to materials optimization and shows faster performance than GP-UCB. Panel b) The number of clusters for the benchmark dataset was initialized to 5 and while this number on average increased during CAMEO's phase mapping, it converged to 5. Demonstrating that improved performance was not due to increased complexity defined by a larger number of clusters.

[0020] FIG. 6. Color coded phase map prior derived from AFLOW.org computed tie-lines (black lines) for the benchmark Fe—Ga—Pd material system.

[0021] FIG. 7. Example Ge—Sb—Te optical data used for phase mapping prior.

[0022] FIG. 8. Panel a) For the 10% material selection out of the 278 materials in the composition spread, the selected 28 materials are indicated with filled circles. Panel b) The order of materials measured during sequential measurement.

[0023] FIG. 9. Modifications made to the Fe—Ga—Pd remnant magnetization as measured as scanning SQUID voltage signals. Panel a) Circles indicate the samples with saturated voltage of 10 V, Panel b) Modified voltage by enhancing main voltage peak at $\mu=Fe_{78}Ga_{16}Pd_6$ and the maximum indicated with a circle.

[0024] FIG. 10. Phase-change temperature mapping of the combinatorial Ge—Sb—Te spread.

[0025] FIG. 11. Structural evolution FCC—Ge—Sb—Te (GST) structure (top) to the Sb—Te structure (bottom) across the line of composition marked in the phase diagram on the right. Peak indices are denoted.

[0026] FIG. 12. The optical bandgap of amorphous (left) and crystalline (right) states for a combinatorial Ge—Sb—Te spread.

[0027] FIG. 13. The performance of the photonic device fabricated by the new nanocomposite PCM, GST467. The symmetric multi-level switching is realized. The inset is the top view of the photonic device used for multi-level switching, endurance test and comparison between $Ge_4Sb_6Te_7$ and $Ge_2Sb_2Te_5$.

[0028] FIG. 14. Comparison of CAMEO, GP-UCB, and random for materials optimization. Each curve shows the percentage of the 100 (post data collection) simulation runs for each active learning scheme which have reached within 1% of the optimal AE_g. Within the first 20 iterations, 31% of CAMEO's runs have reached within 1% of the optimal AE_g, compared to 10% for GP-UBC.

[0029] FIG. 15. Live visualizations for the Fe—Ga—Pd system. Panel a) Samples labeled with the same color are identified to belong to the same phase region. The size of the filled circles indicates probability of estimated region label.

Regions of small circles indicate high likelihood of phase region boundary. Black circles indicate samples that have been measured for x-ray diffraction and target functional property. Red diamond indicates CAMEO's selection of the next sample to measure. Panel b) Color indicates CAMEO prediction of functional property, with each phase region fitted with its own Gaussian process. Here the measured voltage is the signal from scanning SQUID and is proportional to magnetization. Panel c) Color indicates computed GP-UCB acquisition function for (c) Panel d) FMI phase mapping convergence calculated between each subsequent iteration. Panel e) Convergence in maximum functional property value identified computed as the difference in the max identified property in consecutive iterations.

[0030] FIG. 16: The structural information of the nanocomposite phase-change memory material $Ge_4Sb_6Te_7$. The synchrotron radiation measurements show the coexistence of the SbTe impurity and the Ge—Sb—Te matrix.

[0031] FIG. 17: The optical parameters of the $Ge_4Sb_6Te_7$ and its comparison with $Ge_2Sb_2Te_5$. Panel A) The optical parameters of the $Ge_4Sb_6Te_7$ in the amorphous and crystalline states. Panel B) Comparison of the optical contrast indicated by difference in the extinction coefficient k between crystalline and amorphous phases ($k_{crystalline}-k_{a-morphous}$) for the wavelength range of 1000 to 1500 nm.

DETAILED DESCRIPTION

[0032] The foregoing and other aspects of the present disclosure will now be described in more detail with respect to other embodiments described herein. It should be appreciated that the subject matter of the present disclosure can be embodied in different forms and should not be construed as limited to the embodiments set forth herein. Rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the examples provided in the present disclosure to those skilled in the art.

[0033] The terminology used in the description of the examples of the present disclosure is for the purpose of describing particular embodiments only and is not intended to be limiting. As used in the description of the examples and the appended claims, the singular forms "a", "an" and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. Additionally, as used herein, the term "and/or" includes any and all combinations of one or more of the associated listed items and may be abbreviated as "/".

[0034] The term "comprise," as used herein, in addition to its regular meaning, may also include, and, in some embodiments, may specifically refer to the expressions "consist essentially of" and/or "consist of." Thus, the expression "comprise" can also refer to, in some embodiments, the specifically listed elements of that which is claimed and does not include further elements, as well as embodiments in which the specifically listed elements of that which is claimed may and/or does encompass further elements, or embodiments in which the specifically listed elements of that which is claimed may encompass further elements that do not materially affect the basic and novel characteristic(s) of that which is claimed. For example, that which is claimed, such as a composition, formulation, method, system, etc. "comprising" listed elements also encompasses, for example, a composition, formulation, method, kit, etc. "consisting of," i.e., wherein that which is claimed does not

include further elements, and a composition, formulation, method, kit, etc. "consisting essentially of," i.e., wherein that which is claimed may include further elements that do not materially affect the basic and novel characteristic(s) of that which is claimed.

[0035] The term "about" generally refers to a range of numeric values that one skilled in the art would consider equivalent to the recited numeric value or having the same function or result. For example, "about" may refer to a range that is within ±1%, ±2%, ±5%, ±10%, ±15%, or even ±20% of the indicated value, depending upon the numeric values that one skilled in the art would consider equivalent to the recited numeric value or having the same function or result. Furthermore, in some embodiments, a numeric value modified by the term "about" may also include a numeric value that is "exactly" the recited numeric value. In addition, any numeric value presented without modification will be appreciated to include numeric values "about" the recited numeric value, as well as include "exactly" the recited numeric value. Similarly, the term "substantially" means largely, but not wholly, the same form, manner or degree and the particular element will have a range of configurations as one skilled in the art would consider as having the same function or result. When a particular element is expressed as an approximation by use of the term "substantially," it will be understood that the particular element forms another embodiment.

[0036] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one skilled in the art.

[0037] The present disclosure relates to preparation of phase-change materials, such as phase-change memory materials, with improved and enhanced properties. Features include:

1. Phase-change memory materials having improved or enhanced optical bandgap differences between amorphous and crystalline states;

2. Methods of evaluating/examining phase-change memory materials using machine/active learning, such as Bayesian active learning and/or Bayesian manifold learning, to select phase-change memory materials, such as materials having larger optical bandgap differences between phases; and

3. Methods of forming electronic devices, such as photonic switching devices and/or enhanced memory devices, with phase-change memory materials selected using the methods evaluating/examining phase-change memory materials using machine/active learning, such as Bayesian active learning and/or Bayesian manifold learning, as set forth herein.

[0038] Phase-change memory materials include, for example, phase-change memory materials including a material of formula (I):

$$TM_xSb_yTe_z \tag{I}$$

wherein TM is a transition metal, wherein x, y, and z are each independently between about 1-10, between about 3-10, between about 3-9, between about 4-9, between about 3-7 or about 4-7, and wherein the phase-change memory material has an extinction difference between crystalline and amorphous phases (k_c-k_a) greater than about, e.g., 2, in a wavelength range of, e.g., about 1000 nm to about 1500 nm. Other properties of the phase-change memory materials include, for example: a phase-change temperature of, e.g., about 200° C.; a ΔE_g of greater than or equal to about, e.g., 0.76 eV; stability over, e.g., at least about 30,000 cycles,

such as switching between, crystalline and amorphous phases; and, e.g., at least about 15-20 distinct levels, such as, for example, about 17, distinct levels in non-volatile symmetric multi-level switching. It will be appreciated that switching between phases, such as switching between crystalline and amorphous phases, of materials may be accomplished by exposing/subjecting the materials to, for example, light, such as laser light, or an electrical current, but are not limited thereto.

[0039] In some embodiments, phase-change memory materials, such as those materials of formula (I), include a material of formula (II):

$$Ge_xSb_yTe_z \tag{II}$$

wherein x is between about 3.4 and 4, y is between about 5 and 6, and z is between about 6.5 and 7.5, and wherein the phase-change memory material has an extinction difference between crystalline and amorphous phases (k_c-k_a) greater than about, e.g., 2, in a wavelength range of, e.g., about 1000 nm to about 1500 nm. Other properties of the phase-change memory materials include, for example: a phase-change temperature of, e.g., about 200° C.; a ΔE_g of greater than or equal to about, e.g., 0.76 eV; stability over, e.g., at least about 30,000 cycles, such as switching between, crystalline and amorphous phases; and, e.g., at least about 15-20 distinct levels, such as, for example, about 17, distinct levels in non-volatile symmetric multi-level switching, wherein switching between phases, such as switching between crystalline and amorphous phases, of materials may be accomplished by, for example, exposing/subjecting the materials to, for example, light, such as laser light, or an electrical current.

[0040] In some embodiments, systems for analysis, optimization, and/or prediction of characteristics of compositions/materials, and selection of compositions/materials that are suitable for use as, for example, phase-change memory materials, are provided. Systems may include any electrical circuitry that would be appreciated by one skilled in the art configured to perform analysis, optimization, and/or prediction of characteristics of compositions/materials described herein. In some embodiments, circuitry to perform the analysis, optimization, and/or prediction of characteristics of compositions/materials may employ one or more artificial intelligence and/or machine learning methods/techniques, which in some embodiments may be implemented by circuitry configured to perform the artificial intelligence and/or one or more machine learning methods/techniques.

[0041] "Artificial intelligence" is used herein to broadly describe any computationally intelligent systems and methods that can learn knowledge (e.g., based on training data), and use such learned knowledge to adapt its approaches for solving one or more problems, for example, by making inferences based on a received input, such as the received characteristics/properties of compositions/materials having potential use as a phase-change memory material. "Machine learning" generally refers to a sub-field or category of artificial intelligence, and is used herein to broadly describe any algorithms, mathematical models, statistical models, or the like that are implemented in one or more computer systems or circuitry, such as processing circuitry, and which build one or more models based on sample data (or training data) in order to make predictions or decisions related to, for example, analysis, optimization, prediction, and/or selection

of compositions/materials that have potential for use, such as phase-change memory materials, according to embodiments as set forth herein.

[0042] Machine learning techniques utilized may include, for example, neural networks, convolutional neural networks, deep learning, active learning techniques such as Bayesian active learning, Bayesian program learning and/or Bayesian manifold learning, support vector machines, pattern recognition techniques, etc. to analyze, optimize, predict, and/or select characteristics of compositions/materials as set forth herein. Furthermore, the machine learning techniques may implement any one or any combination of the following computational algorithms or techniques: classification, regression, supervised learning, unsupervised learning, feature learning, clustering, decision trees, etc., or the like.

[0043] As one example, an artificial neural network may be utilized by the system to develop, train, or update one or more machine learning models which may be utilized to optimize, predict, and/or determine the characteristics of compositions/materials. An example artificial neural network may include a plurality of interconnected "neurons" which exchange information between each other. The connections have numeric weights that can be tuned based on experience, and thus neural networks are adaptive to inputs and are capable of learning. The "neurons" may be included in a plurality of separate layers which are connected to one another, such as an input layer, a hidden layer, and an output layer. The neural network may be trained by providing training data (e.g., past data related to characteristics/properties of compositions/materials) to the input layer. Through training, the neural network may generate and/or modify the hidden layer, which represents weighted connections mapping the training data provided at the input layer to known output information at the output layer (e.g., classification of characteristics/properties of compositions/materials). Relationships between neurons of the input layer, hidden layer, and output layer, formed through the training process and which may include weight connection relationships, may be stored, for example, as one or more machine learning models within the system, or any circuitry thereof.

[0044] Once the neural network has been sufficiently trained, the neural network may be provided with non-training data (e.g., received characteristics/properties of novel and/or uncharacterized compositions/materials that are not currently part of the training data) at the input layer. Utilizing characteristics/properties knowledge (e.g., as stored in the machine learning model, and which may include, for example, weighted connection information between neurons of the neural network), the neural network may make determinations about the received characteristics/properties of novel and/or uncharacterized compositions/materials at the output layer. For example, the neural network may be used to predict and/or determine if a particular novel and/or uncharacterized composition/material may have characteristics/properties that are suitable for its use as a phase-change memory material.

[0045] Employing one or more computationally intelligent and/or machine learning operations, the system may learn, for example, by developing and/or updating a machine learning algorithm or model based on training data, to predict and/or determine characteristics of compositions/materials based at least in part on knowledge, inferences or the like developed or otherwise learned through training of the system.

[0046] In some embodiments, systems may be operated by and/or on a computer processor, a microprocessor, a microcontroller, or the like, configured to perform the various functions and operations described herein with respect to a system and/or circuitry for analysis of compositions/materials characteristics as set forth herein. In some embodiments, systems and/or electrical circuitry for analysis of compositions/materials characteristics may be configured to execute software instructions stored in any computer-readable storage medium, including, for example, read-only memory (ROM), random access memory (RAM), flash memory, hard disk drive, optical storage device, magnetic storage device, electrically erasable programmable read-only memory (EEPROM), organic storage media, or the like. In some embodiments, analysis of compositions/materials characteristics may include data/information transmission on a communications network without limitation. The communications network may utilize one or more protocols to communicate data/information via one or more physical networks, including local area networks, wireless networks, dedicated lines, intranets, the Internet, and the like.

[0047] Methods used to select compositions/materials, such as phase-change memory materials, accordingly may include, for example, selecting materials using machine learning methods/techniques, such as active learning methods/techniques, for example, Bayesian active learning and/or Bayesian manifold learning, to analyze, optimize and/or predict characteristics, e.g., extinction differences, phase change temperatures, stability, such as cycle stability of switching between crystalline and amorphous phases, of compositions/materials, such as compositions/materials of formula (I) and (II), that may be suitable for phase-change memory materials. The analysis of characteristics may, in some embodiments, be based at least in part on known characteristics that are part of a database for characteristics of compositions/materials, for example, characteristics and/or physical properties of compositions/materials that are available at www.AFLOW.org. In some embodiments, analyzing, optimizing, and/or predicting characteristics of the compositions/materials, as well as selecting compositions/materials, is accomplished via Bayesian active learning and/or Bayesian manifold learning. In some embodiments, the Bayesian active learning and/or Bayesian manifold learning utilized in the methods includes closed-loop autonomous materials exploration and optimization (CAMEO). In some embodiments, methods used to select compositions/materials may include methods as set forth, for example, in Kusne, A. G. et al. On-the-fly closed-loop materials discovery via Bayesian active learning. *Nat. Commun.* 11, 5966 (2020) and supplemental information therefor, the disclosure of which hereby is incorporated by reference in its entirety.

EXAMPLES

[0048] Having described various aspects, the same will be explained in further detail in the following examples, which are included herein for illustrative purposes, and which are not intended to be construed as limiting.

6

Example 1

On-The Fly Closed-Loop Materials Characterization Via Bayesian Active Learning

[0049] Closed-loop autonomous system for materials exploration and optimization (CAMEO) has been demonstrated in real-time control of X-ray diffraction measurement experiments over composition spreads at the synchrotron beamline and in the lab. The algorithm can accelerate phase mapping and materials discovery of a novel solid-state material, with a 10-fold reduction in required experiments, each iteration taking seconds to tens of minutes depending on the experimental task.

Results

[0050] CAMEO uses a materials-specific active-learning campaign that combines the joint objectives of maximizing knowledge of the phase map P(x) with hunting for materials x_* that correspond to property F(x) extrema. Here $x \in \mathbb{R}^d$ is the set of d materials-composition parameters. In particular, subsequent phase mapping measurements are driven by Bayesian graph-based predictions combined with risk minimization-based decision making, ensuring that each measurement maximizes phase map knowledge (see "Methods" sections "M1c-e"). CAMEO accelerates both tasks by exploiting their mutual information via function g (see Eq. 1). Further acceleration is achieved through integration of physics knowledge (e.g., Gibbs phase rule) as well as prior experimental and theory-based knowledge of the target material system.

[0051] These features allow CAMEO to target its search in specific phase regions or to search near phase boundaries where significant changes in the target property are likely to occur, thus exploiting the dependence of materials property on structure. An example of how phase map knowledge can accelerate materials discovery and optimization is shown in FIG. 2a. Off-the-shelf BO methods ignore material structure and assume material properties are a function of only the synthesis parameters, while CAMEO incorporates knowledge that significant changes in properties may occur at phase boundaries. Phase mapping knowledge thus improves property prediction estimate and uncertainty where it matters most.

[0052] For this example a simplified implementation of g is used, switching from phase mapping to materials optimization once phase mapping converges. Materials optimization is focused in the most promising phase region, with a greater importance given to compounds near the phase boundaries. CAMEO thus provides two benefits—a significant reduction of the search space and an improved functional property prediction due to phase map knowledge. Discussion of other approaches to g appears in the "Methods" section. We demonstrate that this physics-informed approach accelerates materials optimization compared to general optimization methodologies that focus on directly charting the high dimensional, complex property function.

$$x_* = \mathrm{argmax}_x[g(f(x), P(x))] \qquad (1)$$

[0053] Here, we explored the Ge—Sb—Te ternary system to identify an optimal phase-change memory (PCM) material for photonic switching devices[30]. PCM materials can be switched between the amorphous and crystalline states with an associated change in resistance and optical contrast which can be accessed on the nanosecond scale or shorter. Various Ge—Sb—Te based PCMs, especially $Ge_2Sb_2Te_5$ (GST225), have been used in DVD-RAM and nonvolatile phase-change random-access memory. We have implemented our strategy for identifying the optimal composition within the ternary for high-performance photonic switching with an eye toward neuromorphic memory applications[31]. Our goal was to find a compound with the highest optical contrast between amorphous and crystalline states in order to realize multi-level optical switching with a high signal-to-noise ratio. The composition range mapped was selected based on the lack of detailed phase distribution and optical property information near known PCM phases. We tasked CAMEO to find the composition with the largest difference in the optical bandgap ΔE_g and hence optical contrast between amorphous and crystalline states. We have discovered a naturally-forming stable epitaxial nanocomposite at a phase boundary between the distorted face-centered cubic Ge—Sb—Te structure (which we refer to as FCC—Ge—Sb—Te or simply GST) phase region and phase coexisting region of GST and Sb—Te whose optical contrast is superior to the well-known GST225 or other compounds within the Ge—Sb—Te ternary. In a direct comparison, a photonic switching device made of the newly discovered material outperforms a device made of GST225 with a significant margin.

[0054] CAMEO satisfies many attributes of a robot scientist, as diagrammed in FIG. 1. The modular algorithm has 'at its fingertips' a collection of information—knowledge of past experiments both physical and computational, materials theory, and measurement-instrumentation science. The algorithm uses this knowledge to make informed decisions on the next experiment to perform in the pursuit of optimizing a materials property and/or maximizing knowledge of a materials system. For example, at each iteration the set of possible phase maps are identified and ranked by Bayesian likelihood, given analysis results of the measured materials. Phase map and functional property likelihoods establish scientific hypotheses and drive further phase mapping and materials optimization and are also presented to the human-in-the-loop who can then (optionally) provide guidance. Specifics are presented in Methods. CAMEO controls lab-based characterization equipment in real-time to orchestrate its own experiments, update its knowledge, and continue its exploration. The more specific implementation of Eq. 1 is shown in FIG. 2.

[0055] CAMEO is based on the fundamental precept that in navigating compositional phase diagrams, enhancement in most functional properties is inherently tied to the presence of particular structural phases and/or phase boundaries. The strategy is, therefore, broadly applicable to a variety of topics with disparate physical properties. The method was first benchmarked and its hyperparameters tuned using a previously characterized composition spread Fe—Ga—Pd, where an entirely different physical property—remnant magnetization, was optimized (see "Methods" for benchmarking method and performance analysis). It was then successfully used to discover a new photonic PCM composition whose ΔE_g (between crystalline and amorphous states) is up to 3 times larger than that of the well-known $Ge_2Sb_2Te_5$.

[0056] For this task, scanning ellipsometry measurements were performed on the spread wafer with films in amorphous (initial) and crystalline states ahead of the CAMEO run, and the raw ellipsometric spectra data were incorporated as a phase-mapping prior. This was performed by

increasing graph edge weights between samples of similar raw ellipsometry spectra during the phase mapping operation (see "Methods" section "M1c"). Thus, the algorithm makes use of information regarding phase distribution across the spread that is "hidden" in the unreduced complex spectroscopic data, which vary non-trivially across the spread. At each iteration, CAMEO identifies the next material to query, indicates the material to the experimentalist (human-in-the-loop) who performs the intensive task of processing the raw optical data to extract AE_g. In parallel, CAMEO remotely controls scanning of the synchrotron beam to collect X-ray diffraction data from the spread wafer with films in the crystalline state. CAMEO first seeks knowledge of the phase map until 80% convergence, and then switches to material optimization (see "Methods" section "M1"). This procedure identified the material with the largest AE_g over 19 iterations taking approximately 10 h, compared to over 90 continuous hours for the full set of 177 composition spots. After data collection CAMEO was bench-marked against common active learning schemes with each run for 100 simulations. CAMEO provides an approximate maximum average 35-iteration lead over the best alternative Gaussian process-upper confidence bounds (GP-UCB) focusing on AE_g optimization in the composition space. The use of the ellipso-metry prior to accelerate phase mapping provides a 25-iteration lead out of the 35. Furthermore, over the 100 runs, CAMEO gets within 1% of the optimal in the first 20 runs 31% of the time compared to GP-UBC's 10% (see FIG. **14**).

[0057] As seen in FIG. **3**, panel b, the optimal composition identified here lies at the boundary between the FCC—Ge—Sb—Te (GST) phase region and the region where there is co-existence of GST and Sb—Te phases. The average composition of the region is Ge4Sb6Te7, and henceforth we refer to the region as GST467. Its AE_g is found to be 0.76±0.03 eV, which is nearly 3 times that of GST225 (0.23±0.03 eV). To investigate the origin of the enhanced AE_g of GST467 at the phase boundary, we have performed high-resolution transmission microscopy of this composition (FIG. **4**, panel a) which revealed a complex nanocomposite structure consisting of GST and Sb—Te phases. As seen in the figure, the phases have grown coherently with the relationship GST Fm-3m (111)//SbTe (001). (see "Methods" for details.)

[0058] This boost in AE_g indeed directly leads to large enhancement in optical contrast as captured in $Ak = k_c - k_a$, the difference in the extinction coefficient (between amorphous (k_a) and crystalline states (k_c)) extracted from the ellipsometry data at different wavelengths (FIG. **4**, panel c). Ak for GST467 is 60-125% larger than that of GST225 in the 1000-1500 nm wavelength range. The superior physical properties of GST467 shown here were reproduced on multiple composition spread wafers.

[0059] We have fabricated photonic switching PCM devices based on the discovered GST467 nanocomposite. With a sequence of laser pulses (energy and pulse width) with varying amplitude sent through the device, the material can be switched between the crystalline and amorphous phases (FIG. **4**, panel b). The device made of the nanocomposite GST467 thin film was found to be stable up to at least 30,000 cycles indicating the high reversibility of the crystallization and quenching processes of the coherent nanocomposite. The one-to-one comparison between the devices fabricated with our GST225 and GST467 films here (FIG. **4**,

panels d and e) shows that GST467 device exhibits a much-enhanced switching contrast resulting in up to 50% more in the number of interval states, important for photonic memory and neuromorphic devices[32,33].

[0060] Recent reports of nanostructured PCM materials, including multilayer and superlattice thin films have highlighted the crucial roles interfaces and defects play in their switching mechanisms leading to faster switching speed and lower switching energies[34,35]. Our finding of GST467 exhibiting significant boost in ΔE_g, and consequently larger optical contrast, underscores the effectiveness of naturally-forming nanocomposites as another approach to enhancing performance of PCM materials, especially for optical switching devices. It is the presence of epitaxial nano-pockets of the SbTe phase in GST467 which is locally modifying the resonant bonding in the GST matrix resulting in the lowered optical bandgap in the crystalline state, which in turn leads to the larger ΔE_g.

[0061] The discovery of a novel PCM material demonstrates that systems similar to CAMEO will fulfill the primary goals of materials design by accelerating the discovery and collection of materials knowledge, streamlining the experiment cycle, improving control over experimental variability, and improving reproducibility, thus improving trust in scientific results. They will also generate reference and benchmark datasets—automatically processed, analyzed, and converted to actionable knowledge with all associated metadata, for developing and improving trust in machine learning tools. Further benefits include automatic knowledge capture to maintain institutional knowledge, maximizing information gain, and reducing use of consumable resources, such as expert time, freeing up experts to work on higher level challenges. Research at the synchrotron exemplifies these resource demands and limitations, where obtaining scientist and equipment time is difficult or expensive. And potentially most impactful, placing labs under the control of AI may greatly reduce the technical expertise needed to perform experiments, resulting in a greater 'democratization' of science[36]. In turn, this may facilitate a more distributed approach to science, as advocated by the materials collaboratory concept[37].

Methods

[0062] M0 Outline of Sections.

[0063] A description of the closed-loop, autonomous system for materials exploration and optimization (CAMEO) scheme can be found in M1 beginning with a detailed description of results and a description of materials and device synthesis and characterization in section M2. The description of CAMEO is broken down into the subsections: M1a detailed results, M1b Initialization and data pre-processing, M1c Phase mapping, M1d Knowledge propagation, M1e Active learning, M1f Statistics and performance measures. The materials and device section is broken down into the subsections: M2a Sample fabrication, M2b Mapping of phase-change temperature, M2c Structural mapping, M2d High-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) of Ge4Sb6Te7 (GST467), M2e Modeling and calculation of the ellipsometry spectra, M2f GST467 photonic device fabrication and measurement.

[0064] M1 CAMEO.

[0065] CAMEO's methodology follows the diagrams of FIGS. **1** and **2**c (main text), where active learning drives measurement and expert input. Active-learning-driven syn-

thesis and simulation are excluded for this work. The materials are pre-synthesized as a composition spread and, if desired, the AFLOW.org density functional theory (DFT) simulations are run prior to CAMEO's control of the X-ray diffraction measurement. For this work, DFT calculations were run to build a prior for the Fe—Ga—Pd phase mapping. DFT data were not used for the Ge—Sb—Te material system, instead ellipsometry data was used to build a phase mapping prior. The combinatorial library is physically loaded into the high-throughput X-ray diffraction system, and any data captured from external or internal databases is automatically imported into CAMEO. All preliminary data is analyzed to build the first estimated phase map along with uncertainty quantifications. This kicks off the iterative autonomous process where the phase map and material property estimate and estimate uncertainty are used to inform the active-learning-driven selection of the material to query next. At each iteration, CAMEO selects a material to study and requests and obtains structure and functional property data for the query material, with automatic X-ray diffraction pre-processing. In parallel, results and predictions are presented to the expert user and pertinent knowledge is captured from the expert. All gathered knowledge is then stored in a database. A description of the capabilities of CAMEO are presented in Table 1.

optimization). However, g is a general function, flexible for other implementations, such as combining knowledge of composition, lattice structure, and functional property (as well as microstructure, and other information) in one integrated acquisition function. Future work will explore the pros and cons of these other implementations.

[0068] For this work, the first set of iterations maximize phase map knowledge until the estimated phase map convergences to the user defined threshold c, at which point the system switches to materials property optimization. A separate Gaussian Process is fit to each individual phase region for the functional property, allowing for phase region dependent hyperparameter optimization. This exploits the CSP relationship to improve functional property prediction accuracy, accelerate materials optimization, and provide potential computational resource savings. The phase regions are then ranked by the maximum expected functional property value and the top R regions are selected for optimization, with R a user defined variable. Here R is set to 1. Optimization balances exploitation and exploration through the mean $\mu(x_r)$ and weighted variance $\beta\sigma(x_r)$ (the iteration dependent β follows ref. 38 and is described below). The optimization acquisition function also allows the user to target points closer or further from phase boundaries via $\gamma d(x_r)$, where $d(x_r)$ is the distance from point x_r to the nearest

TABLE 1

Knowledge and Control implemented in CAMEO	
Knowledge and Control	CAMEO
Knowledge: Past experiments both physical and computational	Automated access to experimental and density functional theory materials structure databases. Includes Inorganic Crystal Structure Database and AFLOW.org.
Knowledge: Materials physics theory	Phase mapping and structure theory including Gibbs phase rule via constraint programming
Knowledge: Materials synthesis and processing	NA
Knowledge: Measurement science	X-ray diffraction simulation capability using structure data as input
Control: Synthesis control	NA
Control: Characterization	X-ray diffraction: high-throughput X-ray diffraction system[1] at the Stanford Synchrotron Radiation Lightsource (SSRL) and Bruker D-8*
Control: Communication	GUI for user interface; Interface to databases to store and share knowledge with experts and other AIs; Network interface for instrument control

[0066] CAMEO's specific implementation of Eq. 1 is shown in Eq. 2:

$$g(x) = \begin{cases} P(x), & c < 80\% \\ F(x_r) = \mu(x_r) + \beta\sigma(x_r) + \gamma d(bfx_r) & \text{else} \end{cases} \quad (2)$$

[0067] Recent use of active learning in materials science seeks to optimize functional material properties as a function of only the material synthesis parameters, e.g., composition. However, a material's properties are clearly not just dependent on its composition. Fundamental to functional properties are a material's lattice structure, microstructure, stress, etc. The general function g provides a framework for incorporating these different pieces of knowledge as well as an active learning scheme for acting on the knowledge. Here we take the simplest action scheme, switching between one phase of optimization (phase mapping) to another (materials

phase boundary and γ is a user-defined parameter—negative (positive) to emphasize points near the edge (center) of the phase region. Here the value is set to 10. Myopia to particular phase regions can be removed with an additional exploration policy.

[0069] Pre-synthesized (pseudo) ternary combinatorial spreads are used to provide a pool of hundreds of materials to investigate. While for this demonstration the autonomous system must select samples from the given pool of pre-synthesized samples, this is only a limit of the current physical experimental system and not a limit of the presented ML methodology.

[0070] M1a CAMEO Detailed Results.

[0071] CAMEO was benchmarked on a material system previously studied[39]. Efficacy was compared to a range of alternative methods as shown in FIG. 5 with phase mapping performance measured with the Fowlkes-Mallows index (FMI) and Bayesian optimization performance measured by

minimum percent deviation from optimal. The mean performance and 95% confidence intervals over 100 iterations are plotted in FIG. **5**. The algorithm provides significant accuracy improvement and lower variance in phase mapping. Additionally, each level of increased physical knowledge further accelerates phase mapping. The benchmark optimization challenge was to maximize a functional property that is a simple function of composition with one broad, dominant peak in one phase region and a smaller peak with a maximum in another phase region. For this simple challenge, CAMEO provides improved performance compared to the next best optimization scheme—Gaussian process-upper confidence bounds (GP-UCB). For more information about the benchmarking process see (Active learning—materials optimization: benchmark system).

[0072] Once tuned, CAMEO was placed in active control over the high-throughput X-ray diffraction system at SLAC and a commercial in-house diffraction system. Here, the material optimization goal was to identify an optimal phase change material in the Ge—Sb—Te system, characterized by maximizing ΔE_g—the difference between the amorphous and crystalline optical bandgap. Scanning ellipsometry measurements were performed on the spread wafer in amorphous and crystalline states ahead of the CAMEO run, and we fed the unprocessed ellipsometric spectra as a prior for building the phase map model. At each iteration, the query material was indicated to the experimentalist (human-in-the-loop) who then performed the intensive task of processing the raw optical data to obtain ΔE_g and provided this data to CAMEO (see section "M2e" for full description). This procedure identified the material with the largest ΔE_g over 19 iterations taking approximately 10 h, compared to 90 h for the full set of 177 materials. A post-analysis is shown in FIG. **4**, where 100 runs are performed comparing CAMEO to alternative methods. CAMEO provides an approximate maximum average 35-iteration lead over GP-UCB. More importantly, the algorithm is able to mine and make use of information regarding phase distribution across the spread hidden in the complex raw spectroscopic data.

[0073] M1b System Initialization and Data Pre-Processing

[0074] Physical system initialization. The system is initialized by loading the composition spread into the X-ray diffraction system, either the Bruker D8* or the SSRL diffraction synchrotron beamline endstation. For the SSRL system, a network connection is used for sending commands to the X-ray diffraction system via the SPEC interface[40]. Exposure time for each point measurement was 15 s.

[0075] Importing external data: ICSD and AFLOW.org. The user first indicates the material system of interest. A database of known stable phases, derived from past phase diagrams, is then used to automatically identify pertinent phases. Structure data is then automatically assembled for these phases from the Inorganic Crystal Structure Database (ICSD)—a database of critically evaluated experimental structures, and the AFLOW.org[41] density functional theory database. All retrieved structures are then used to generate simulated diffraction patterns through a call to Bruker's Topas[42]. After data is collected from the databases, the pool of material samples is updated to contain both the samples on the composition spread and those derived from databases. Previously it was shown that external structure data improved phase mapping performance in the case of exhaustive data collection[42]. For this work, the AFLOW.org computed ternary energy hull is imported and converted to

region labels which are used as phase region (i.e. cluster) priors, see FIG. **6** and Phase mapping prior.

[0076] Initialize phase mapping. Phase mapping is initialized with a user-selected expected number of phase regions for the material system, 5 for Fe—Ga—Pd and 10 for Ge—Sb—Te. While this number is used to initialize the phase map model, the phase mapping technique will converge to either a larger or smaller number of phase regions as described in the GRENDEL (graph-based endmember extraction and labeling) section. All other phase mapping hyperparameters were optimized on the benchmark system, and these values were used without modification for the Ge—Sb—Te system. Other default parameters include: graph distance multiplier is 1.2 and max number of iterations is 100.

[0077] Selection of first sample to seed processes. If prior material structure data is imported, such as data from AFLOW.org, that knowledge is used to initialize phase mapping (see Phase mapping prior), with the active learning criterion used to select the most informative material to query next. However, if no such prior data is used, the first sample queried can be selected randomly or using some other informative process. For benchmarking, the initial material was selected at random with uniform probability. For the live application to the Ge—Sb—Te system, the first sample was selected to be the one at the composition center of the materials on the composition spread. This sample was selected as it is potentially the most informative, given no other knowledge of the samples. The live run for the Ge—Sb—Te system completed after all the materials were measured, allowing for later analysis of active learning methodologies. To compare these methods, the initial material was again selected at random with a uniform prior.

[0078] Measurement and data pre-processing: collection, integration, background subtraction. Once the next query material has been identified, the system then measures the query material for X-ray diffraction using a programmatically generated script via SPEC for the SLAC high-throughput system or a GADDS script for the Bruker system. For the Bruker system, the diffraction image is integrated into a diffraction pattern automatically, and for the SLAC system, integration is performed as well[43]. The background signal is then automatically identified and subtracted.

[0079] The background signal from sample to sample can vary significantly, requiring a background subtraction method capable of handling these variations. For both the SLAC and Bruker diffraction measurements, Matlab's envelope function with the 'peak' setting and the parameter value of 50 was used to identify and remove the background curve.

[0080] M1c Phase mapping

[0081] Main method: GRENDEL—list of physical constraints. Phase mapping was performed using the physics-informed phase region and constituent phase identification method GRENDEL[44]. This method represents the material composition space as a graph, where each material is represented by a vertex that is connected by edges to neighboring materials in composition space (or wafer x-y coordinate). Neighboring materials are defined by Voronoi tessellating the composition space[44]. Mathematically, $G=\{V, E\}$, where V is the set of vertices, E is the set of edges with all edge weights set to 1. G is used to define a Markov Random Field (MRF)[45] where materials identified with the same vertex label belong to the same phase region, and each

phase region is described by a set of constituent phases. This method encodes a list of physical constraints through the methods listed in Table 2.

a prior for phase mapping. For materials that share a graph edge and a clustering label, the edge weight in E_p is set to 1, otherwise the edge connecting them is removed from E_p.

TABLE 2

List of physical constraints in [(32)] method and associated encoding methods.

Physical Constraint	Encoding Method
Phase regions are cohesive and phase boundaries are continuous	1. If two or more set of vertices share the same phase region label but are not connected by vertex neighbors, differing labels are assigned to the disconnected sets. 2. The Markov Random Field smoothness constraint
Materials of similar synthesis and processing parameters have similar properties	1. Markov Random Field smoothness constraint 2. Harmonic Energy Minimization for label propagation
Abundances of phases is non-negative	Karush-Kuhn-Tucker conditions[2]
X-ray diffraction intensity is non-negative	Karush-Kuhn-Tucker conditions[2]
Soft Gibbs Phase Rule-Upper bound limit on number of constituent phases	Upper limit on number of endmember limits allowed in each phase region
Identified endmembers should be physically realizable	Volume constraint on identified/predicted endmembers

[0082] This method identifies a phase map for hundreds of samples in tens of seconds, on the same order of X-ray diffraction measurements at SSRL which typically takes 30 s, and measurements on the Bruker D8 which takes over 10 min.

[0083] GRENDEL hyperparameters include the MRF segmentation (i.e., graph cut) weight and the balance between the material-phase region matrix based on clustering and that based on phase mixture[44]. As the graph cut weight is increased, a greater number of clusters becomes possible, increasing the phase mapping performance using the measured described in the text, while also increasing cluster complexity. For the Fe—Ga—Pd a graph cut weight of $w_{gc}=100*n^3/_5{}^3$ was found to output the desired number of clusters n. The full set of phase mapping parameters are described in the text.

[0084] During the GRENDEL process, if the number of clusters drops below 90% of the number of clusters used when starting the process, GRENDEL is terminated and the computed phase map labels and constituent phases from the previous internal GRENDEL iteration are output.

[0085] Phase mapping prior. Material property data is incorporated into the MRF model as a prior through the edge weights of the composition graph G, where the original edge weights of G are modified by a functional property graph G_p with edge weights of 0 (disconnected) or 1 (connected) and f: E, E_p->E. If e E En E_p then e=1+ϵ_e else e=1−ϵ_e. The value of ϵ_e was varied for the benchmark material system and the value of ϵ_e=0.5 selected as it demonstrates clear improved phase mapping performance during the first active learning selected measurements and worse performance near the end of the run. This is to be expected as prior knowledge can benefit initial analysis but can overwhelm knowledge gained from data if the prior is weighted too heavily. A smaller (larger) value of ϵ_e demonstrates a smaller (larger), similar effect.

[0086] For the benchmark system an AFLOW.org based phase map prior was used, as shown in FIG. 6, where the AFLOW.org tie-lines are used to define regions. Points that fall in the same region are given the same label, resulting in

[0087] For the Ge—Sb—Te material system, the prior was determined based on optical data collected. For each material, the complex reflectance ratio amplitude ψ and phase difference Δ for the amorphous and crystalline phases were collected for the set of angles e={50°, 55°, 60°, 65°, 70°} relative to the laser's plane of incidence, creating 20 spectral measurements for each material consisting of different measurement types $m \in \{\psi^{crystalline}, \psi^{amorphous}, \Delta^{crystalline}, \Delta^{amorphous}\}$. Example optical data used for the prior is shown in FIG. 7.

[0088] To define a prior for the phase diagram, the set of all spectra are reduced into a set of similarity weights defining a similarity of 0 or 1 for each pair of samples, which can then be used to evaluate E_p. The following equations are used for mapping of spectra to similarity values. First the Euclidean difference between each set of materials (i, j) is computed for each spectral measurement type and angle {m, e}. These differences are then averaged for each pair of samples (i, j) over the set of angles e and then normalized to between 0 and 1 for each measurement type m. These values are then averaged again over measurements m, resulting in a final dissimilarity value for each pair (i, j). A threshold is then used to convert the continuous dissimilarity values to 0 or 1, defining whether an edge between (i, j) exists (1) or does not (0). The threshold of $D_{Threshold}$=0.07 was selected to achieve a ratio of |E_p|/|E|=0.49≈0.5, i.e., the prior removes approximately half the edges from the initial graph.

$$D_{i,j}^m = \underset{\theta}{\text{mean}} \ [d_{Euclidean}(m_{i,\theta}, m_{j,\theta})] \quad (3)$$

$$D_{i,j}^{\prime m} = [D_{i,j}^m - \min D_{i,j}^m]/[\max D_{i,j}^m - \min D_{i,j}^m] \quad (4)$$

$$\overline{D}_{i,j} = \underset{m}{\text{mean}} \ [D_{i,j}^{\prime m}] \quad (5)$$

$$G_p = \overline{D}_{i,j} < D_{Threshold} \quad (6)$$

[0089] M1d Knowledge Propagation

[0090] Phase mapping knowledge propagation. Once the phase map has been identified for the given data, the phase region labels must be propagated to the materials that have

yet to be measured for structure. To exploit the graph-described data manifold, the semi-supervised learning technique Gaussian random field harmonic energy minimization[46] (HEM) is used. HEM computes the likelihood of each material belonging to each phase region and then assigns each material to the phase region with the greatest likelihood, thus defining the most likely phase map for the full set of materials on the composition spread. Using this information, alternative phase maps can also be identified along with their likelihoods. The edge weights E_p define the similarity matrix used to define the graph Laplacian.

[0091] Phase mapping knowledge propagation—comparison method: nearest neighbor (NN). The phase mapping knowledge propagation harmonic energy minimization method is compared to the use of 1-nearest neighbor, where any material without a phase region label takes on the label of its 1st nearest neighbor with a label. First nearest neighbor was implemented using MATLAB's knnsearch function with default parameters.

Functional property knowledge propagation: Gaussian process regression. GPR was implemented using MATLAB's 'fitrgp' function with default parameters.

[0092] M1e Active Learning.

[0093] In the Bayesian optimization[47] literature, the following formalism is used:

$$y=f(x)+\varepsilon \tag{7}$$

$$x_*=\mathrm{argmax}_x(f(x)) \tag{8}$$

where y is the target property to be maximized, $x \in \mathbb{R}^d$ is the set of material synthesis and processing parameters to be searched over, f(x) is the function to optimize, ε is typically independent stochastic measurement noise, and x_* defines the material synthesis and processing parameters that result in the maximal material property (for the minimum, replace argmax with argmin). When f(x) is unknown, a surrogate model is used to approximate it based on given data. The surrogate function is then used to identify the best next material to study. Each subsequent material is selected to identify the optimal material x_* in the smallest number of experiments possible. Identifying extrema of a function involves a balance between exploiting prior data to identify nearby extrema and exploring to identify extrema far from prior data. An alternative active learning objective is to select experiments that will best improve the overall prediction accuracy of the surrogate model, or in other words, select experiments to most efficiently learn the unknown function f(x). Such a campaign learns the general trends of f(x), which is highly useful when attempting to quantify anomalous behavior of novel materials.

[0094] Active learning for phase mapping: risk minimization. The active learning method used to select the next material to query for phase mapping is based on risk minimization[46]. HEM propagates phase region labels to unmeasured material and identifies the likelihood of each material belonging to each phase region. These likelihoods can be aggregated to define the set of potential phase diagrams and their associated likelihoods. The set of potential phase diagrams form a hypothesis space of phase diagrams. Risk minimization seeks to identify the optimal material to query next for its structure that will most rapidly whittle down the hypothesis set and most rapidly hone-in on the optimal phase map for the full set of materials, i.e.,

minimize expected total phase region label misclassification error and equivalently maximize knowledge of the phase map.

[0095] Active learning comparison methods: random, sequential, and 10% sampling. The risk minimization method is compared to (1) random sampling—selecting each subsequent material at random from the wafer, with a uniform prior, (2) sequential sampling—where each sample is selected in the order it appears on the wafer, and (3) where 10% of the materials are selected in a pre-determined design. Random sampling is expected to provide increasingly poor performance relative to active learning as the search space increases in dimension due to the curse of dimensionality[48]. The pre-determined 10% selection of materials in (3) are chosen to provide maximal coverage of the composition space. However, the use of 10% is not a generalizable benchmark. For a given number of data points, the density of data points decreases as the dimensionality of the composition space increases, with each point describing a larger volume. The optimal number of benchmark materials is thus dependent on the expected size of phase regions. If smaller phase regions are expected, a larger number of materials will be required to identify the phase regions.

[0096] The Fe—Ga—Pd composition spread contains 278 samples. For the 10% sampling, the 28 samples are indicated in FIG. 8, panel a. They were selected to provide uniform coverage of the composition space described by the spread. For the sequential sampling, the order of samples is shown in FIG. 8, panel b.

[0097] Active learning—materials optimization: Gaussian process upper confidence bounds. For CAMEO and GP-UCB the iteration dependent weight parameter β is used[38].

$$\beta=2 \log(DI^2\pi^2/6\lambda) \tag{9}$$

Where D is the total number of samples, I is the current iteration number, and $\lambda=0.1$.

[0098] Active learning—CAMEO: phase mapping convergence. The phase maps identified at each iteration i is compared to the iteration (i-4) using the FMI performance measure. Convergence is defined as FMI>=80%.

[0099] Active learning—materials optimization: benchmark system. The target optimization for the benchmark system is maximizing remnant magnetization, as measured by scanning SQUID voltage. One modification was made to the remnant magnetization signal: The signal saturates over a large range of the composition spread. For BO benchmarking, it is preferred that one material is identified as the optimum. As such, the saturated values were modified with a squared exponential function, in effect "hallucinating" the remnant magnetization values as if sensor saturation had not occurred, converting the signal from FIG. 9, panel a to FIG. 9, panel b. The squared exponential function used to modify the voltage was defined in cartesian space. For the ternary composition (a_{Fe}, b_{Ga}, c_{Pd}):

$$x=(b/100)+(c/100)*\sin(30°), y=(c/100)*\sin(60°) \tag{10}$$

$$N_{xy}(\mu=(0.19,0.05),\sigma^2=0.001) \tag{11}$$

[0100] Mapping to ternary space gives $\mu=Fe_{78}Ga_{16}Pd_6$.

[0101] M1f Statistics and Performance Metrics

[0102] Confidence interval. The 95% confidence interval was computed for the variable of interest over 100 experiments at the given iteration with:

$$CI_{95} = \left(\frac{\sigma}{\sqrt{n}}\right)F^{-1}(p, v) \tag{12}$$

Where F^{-1} is the inverse of the Student's t cumulative distribution function, σ is the standard deviation, $n=100$ is the number of experiments, $p=\{2.5\%, 97.5\%\}$, and $v=99$ is the degrees of freedom.

[0103] Phase mapping. Phase mapping performance is evaluated by comparing phase region labels determined by experts with those estimated by CAMEO for the entire phase map (after the knowledge propagation step). To evaluate system performance, the Fowlkes-Mallows Index (FMI) is used, which compares two sets of cluster labels. The equations are presented below for the expert labels $l \in L$ and the ML estimated labels $\bar{l} \in \bar{L}$, where the labels are enumerated $L \rightarrow \mathbb{N}$ and $\bar{L} \rightarrow \mathbb{N}$.

[0104] If the number of phase regions is taken to be too large by either the user or the ML algorithm while the phase mapping is correct, some phase regions will be segmented into sub-regions with the dominant phase boundaries preserved. For example, peak shifting can induce phase region segmentation[42]. To ensure that the performance measures ignore such sub-region segmentation, each estimated phase region is assigned to the expert labeled phase region that shares the greatest number of samples. The number of phase regions is monitored to ensure that increases in model accuracy are not driven by increases in model complexity.

$$\text{Fowlkes-Mallows Index: } FMI = \frac{TP}{\sqrt{(TP+FP)(TP+FN)}}$$

$$TP = \frac{1}{2}\Sigma_i\Sigma_j(l_i = l_j \& \hat{l}_i = \hat{l}_j) \tag{14}$$

$$FP = \frac{1}{2}\Sigma_i\Sigma_j(l_i \neq l_j \& \hat{l}_i = \hat{l}_j) \tag{15}$$

$$FN = \frac{1}{2}\Sigma_i\Sigma_j(l_i = l_j \& \hat{l}_i \neq \hat{l}_j) \tag{16}$$

$$TN = \frac{1}{2}\Sigma_i\Sigma_j(l_i \neq l_j \& \hat{l}_i \neq \hat{l}_j) \tag{17}$$

[0105] M1g Live Visualizations.

[0106] CAMEO provides live visualizations to support algorithm interpretability. Phase mapping is supported by FIG. 15, panels a and b, where each filled circle indicates a sample in the combinatorial library. Samples labeled with the same color are identified as likely belonging to the same phase region. Samples with a black border have been measured for X-ray diffraction and the sample indicated by the red diamond is the sample that CAMEO will query next. The size of each sample marker indicates the probability of the sample belonging to the color-indicated phase region. Thus, areas of the composition diagram where filled circles are small indicate a high likelihood region for a phase boundary. FIG. 15, panel b presents the CAMEO predicted values of the target functional property. Here each predicted phase region is analyzed using its own Gaussian process (see "Methods" Section "M1"). FIG. 15, panel c shows the associated Gaussian process predicted variance which has been mapped using the sigmoid function to values between 0 and 1.

[0107] FIG. 15, panel d presents the FMI score between each subsequent phase mapping iteration, quantifying the percent change in phase mapping from iteration to iteration. FIG. 15, panel e shows convergence in materials property optimization computed as the difference in the max identified property in consecutive iterations.

[0108] Bayesian optimization. Bayesian optimization performance is measured with minimum percent deviation from optimal, related to simple regret.

$$\text{Minimum percent deviation from optimal} = 100\% * \min_i \frac{(p_{target} - p_i)}{p_{target}} \tag{18}$$

$$\text{Simple Regret} = \min_i(p_{target} - p_i) \tag{19}$$

[0109] M2 Materials Synthesis and Characterization

[0110] M2a Sample fabrication. Amorphous thin-film composition spreads encompassing a region of the Ge—Sb—Te ternary (separated into 177 samples using a gridded physical shadow mask) were fabricated on 3-inch silicon wafers with SiO_2 layers (2 μm) by co-sputtering Ge, Sb, and Te targets at room temperature. Different (average) thickness composition spreads (covering the same composition range) were fabricated for different measurements: they were 20 nm, 100 nm, 200 nm, and 500 nm for optical, structural, resistance, and composition mapping, respectively. To obtain a crystalline state, some of the wafers were annealed at 300° C. for 10 min following their characterization in the amorphous state.

[0111] The composition mapping of the spreads is measured using the wavelength dispersion spectroscopy. For every separated sample region on a spread, three random spots are measured, and the average composition value is used for the actual stoichiometry mapping in FIG. 10.

[0112] M2b Mapping of phase-change temperature. Upon increasing the temperature, a phase-change memory material undergoes a structural transition from amorphous to crystalline states with up to four orders of magnitude in the change of resistance. The temperature at which the resistance drop takes place can be taken as the phase-change temperature, T_{cp} (FIG. 10). The entire spreads were measured in a scanning four-probe station combined with a Keithley 2400 from room temperature up to 300° C. T_{cp} of GST467 was found to be ~200° C., which is much higher than that of GST225 (≈140° C.). The higher T_{cp} of GST467 indicates higher stability of the amorphous state of GST467 compared to GST225.

[0113] M2c Structural mapping. Synchrotron diffraction on crystallized spreads was carried out at Beamline 10-2 at SLAC. In addition to the remote-controlled CAMEO run, we have also carried out diffraction of entire spreads in order to obtain the complete structural phase mapping of the probed Ge—Sb—Te region and to verify the accuracy of the phase diagram determined by CAMEO. FIG. 11 shows an example set of diffraction patterns taken across the spread. Along the marked line in the composition map, the evolution of diffraction patterns indicates phases going from the distorted FCC—Ge—Sb—Te (GST) structure region to the phase co-existence region (GST and Sb—Te) to the Sb—Te region.

[0114] SbTe ($R\bar{3}m$), Sb2Te2 ($P\bar{3}m$), and Sb2Te3 ($R\bar{3}m$) all have very similar diffraction patterns and atomic projections of the [100] zone-axis, except for different lattice periods along the [001] direction. These three phases are present across the Sb—Te region depending on the local composition on our spread. The predominant Sb—Te phase in GST467 is SbTe (below).

[0115] M2d HAADF-STEM of GST467. We have performed cross-sectional high-angle annular dark-field scan-

ning transmission electron microscopy (HAADF-STEM) measurements on the GST467 thin film and found that there are nanometer-sized SbTe regions grown coherently inside the distorted cubic GST matrix as shown in FIG. **4**a of the main text. To distinguish between similar phases (SbTe, Sb2Te2, and Sb2Te3), analysis of electron diffraction rings was carried out (not shown here), and Sb—Te phase in the GST467 was identified to be Sb_1Te_1.

[0116] M2e Modeling and calculation of the ellipsometry spectra. The experimental ellipsometry data (J.A. Woollam company) of the combinatorial Gs—Sb—Te spread was analyzed in the range from 200 to 1000 nm using the CompleteEASE software. The dielectric function e(ω) used in the model contains[49] (1) a constant, (2) a Drude-type contribution for free carriers in the case of crystalline state, and (3) a Tauc-Lorentz oscillator to describe the onset of optical transition:

$$\text{amorphous state:} \varepsilon(\omega)=\varepsilon_{const}+\varepsilon_{Tauc\text{-}Lorentz}(\omega), \quad (20)$$

$$\text{crystalline state:} \varepsilon(\omega)=\varepsilon_{const}+\varepsilon_{Drude}(\omega)+\varepsilon_{Tauc\text{-}Lorentz}(\omega) \quad (21)$$

[0117] For the Drude model:

$$\varepsilon_{Drude}(\omega) = \varepsilon_1(\omega) + i \cdot \varepsilon_2(\omega) = 1(\varepsilon(\infty)) - \frac{\omega_p^2}{\omega^2 + \Gamma^2} + \frac{\omega_p^2 \cdot \Gamma}{\omega \cdot (\omega^2 + \Gamma^2)} \quad (22)$$

where $\omega_p = \sqrt{N \cdot e^2/m \cdot \varepsilon_0}$, and ω_p is the plasma frequency, Γ is the collision frequency.

[0118] For the Tauc-Lorentz model:

$$\varepsilon_{Tauc\text{-}Lorentz}(\omega) = \varepsilon_1 + i \cdot \varepsilon_2(\omega) = \quad (23)$$

$$\varepsilon_1(\infty) + \frac{2}{\pi}P \int_{E_g}^{\infty} \frac{\xi \varepsilon_2(E)}{\xi^2 - E^2} d\xi + i \cdot \frac{AE_0 \Gamma(E - E_g)^2}{E\left[(E^2 - E_o^2)^2 + \Gamma^2 E^2\right]} \Theta(E - E_g)$$

Where A is the prefactor, E_o is the peak in the joint density of the state, E_g is the optical bandgap, and Γ is the broadening parameter.

[0119] The optical parameters, e.g., refractive index n and extinction coefficient k, as well as the optical bandgap, can be extracted from these models. In order to confirm the accuracy of the optical parameters, one needs to check if the fitting curves as well as a set of fitting parameters, e.g. thickness, carrier density, and surface roughness, can be used to analyze the experimental data of the samples for the entire spread wafer. Typically, a fitting procedure requires repeated steps in order to fine-tune the parameters manually to optimize the results, and some samples require more manual fitting steps for setting the range and the starting values of the parameters than others.

[0120] For the CAMEO run, the unprocessed raw ellipsometry data taken at each composition spot (for crystalline and amorphous states) are used as the prior (see section "M1c, Phase mapping prior"). Once a spot is identified as a possible composition with enhanced AE_g (the difference in the optical bandgap between the amorphous and crystalline state), the fitting procedure above is carried out on the raw data, and the value of AE_g is computed, the process of which includes manual inspection of fitting parameters. Depending on the number of repeated steps, each computation can take

up to 20 min. at a composition spot, and the AE_g value is then fed back to the CAMEO algorithm.

[0121] The complete mapping of the optical bandgap of amorphous and crystalline states measured and calculated from one spread are shown in FIG. **12**. In the amorphous state, the Ge—Sb—Te based compounds are effectively covalently-bonded semiconductors with large optical band-gaps[50,51]. With changing composition, there is variation in bonding leading to slight shift in the optical bandgap shown in FIG. **12**. In the crystalline state, the resonantly-bonded p orbitals can delocalize the carriers resulting in the reduced bandgap[49,52-54], leading to the large contrast between the amorphous and crystalline states. In the distorted cubic phase (i.e., GST), with changing composition, the local distortion (i.e., Peierls distortion) due to vacancies[52,55,56] would modify the resonant bonding shifting the optical bandgap. In the Sb—Te phase, the optical bandgap also varies with the changing composition in FIG. **12**. When the epitaxial nanocomposite with the SbTe phase are coherently and homogeneously grown in the GST matrix as shown in FIG. **4**a, the nano SbTe phase can act as the impurity dopant phase in the GST matrix.

[0122] M2f $Ge_4Sb_6Te_7$ photonic device fabrication and measurement. Photonic switching devices were fabricated out of GST467 films (FIG. **13**). The 30 nm thick nanocomposite GST467 thin film was sputtered on a 330 nm thick Si_3N_4 layer on an oxidized silicon wafer. A 10 nm thick SiO2 protection layer was then coated on the top of the GST467 thin film. Using e-beam lithography and inductively coupled plasma etching, a 1.2 μm wide photonic waveguide was fabricated. Then the GST467 thin film was patterned into disk shaped features 500 nm in diameter on the top of the waveguide, and they were encapsulated with a 200 nm thick Al2O3 layer as shown in the inset of FIG. **13**.

[0123] A symmetric multi-level switching of the photonic device was investigated as shown in FIG. **13**. In order to provide and precisely control the pump pulses to quench or anneal the GST467 thin film in steps, pulses from a CW pump laser were first modulated by an electro-optic modulator and then sent into an erbium-doped fiber amplifier followed by a variable optical attenuator. The output of the optical signal was collected with a photodetector. During the annealing process, a sequence of pump pulse (50 ns, 2 mW) train was applied to the photonic device. In the quenching process, a sequence of 50 ns pump pulses with gradually increased amplitude was sent into the waveguide.

Example 2

Properties of a Novel Nanocomposite Phase-Change Memory Material Characterized Via Bayesian Active Learning

Fundamental Properties of the Novel Nanocomposite Phase-Change Memory Material

[0124]

Fundamental property	Features	Checking method
Composition	$Ge_4Sb_6Te_7$	Wavelength dispersion spectroscopy

-continued

Fundamental property	Features	Checking method
Phase-change temperature	200° C.	Resistance-temperature mapping [1]
Structure	The SbTe impurity dopants intergrow with the distorted cubic Ge-Sb-Te matrix	Synchrotron radiation and STEM measurements [2]
Optical property	Improvement in the optical contrast between the amorphous and crystalline states	Ellipsometry spectra [3]

[0125] [1] Upon increasing the temperature, a phase-change memory material undergoes a structural transition from amorphous to crystalline states with up to four orders of magnitude in the change of resistance. The temperature at which the resistance drop takes place can be taken as the phase-change temperature, T_{cp}. This new nanocomposite phase-change memory material was checked in a scanning four-probe station combined with a Keithley 2400 from room temperature up to 300° C. The phase-change temperature of $Ge_4Sb_6Te_7$ was confirmed to be approximately 200° C., which is much larger than that of well-known $Ge_2Sb_2Te_5$.

[0126] [2] Synchrotron diffraction on crystallized $Ge_4Sb_6Te_7$ was carried out at Beamline 10-2 at SLAC. A weak SbTe diffraction peak was observed in FIG. 16. A cross-sectional HAADF-STEM measurement were taken on the $Ge_4Sb_6Te_7$ thin film and the nanometer-sized SbTe regions grows coherently inside the distorted cubic Ge—Sb—Te matrix (see, FIG. 4, panel a).

[0127] [3] The ellipsometry measurements were performed on $Ge_4Sb_6Te_7$ on in amorphous and crystalline states to analyze the optical parameters, e.g., refractive index n (FIG. 17, Panel A, left) and extinction coefficient k (FIG. 17, Panel A, right). Compared with the widely-used $Ge_2Sb_2Te_5$, the contrast Δk between the amorphous and crystalline states ($\Delta k = k_{crystalline} - k_{amorphous}$) are largely enhanced, e.g., two times larger than that of $Ge_2Sb_2Te_5$ at the wavelength around 1500 nm as shown in FIG. 17, Panel B.

Advanced Performance of the Novel Material in Photonic Devices[4]:

[0128]

Performance	Features
Endurance	Stable up to at least 30000 cycles
Multi-level switching [5]	Non-volatile and symmetric multi-level switching Up to 50% more in the number of interval states than $Ge_2Sb_2Te_5$

[0129] [4] Device fabrication: photonic switching devices were fabricated out of $Ge_4Sb_6Te_7$ films. The 30 nm thick nanocomposite $Ge_4Sb_6Te_7$ thin film was sputtered on a 330 nm thick Si_3N_4 layer on an oxidized silicon wafer. A 10 nm thick SiO_2 protection layer was then coated on the top of the $Ge_4Sb_6Te_7$ thin film. Using e-beam lithography and inductively coupled plasma etching, a 1.2 μm wide photonic waveguide was fabricated. Then the $Ge_4Sb_6Te_7$ thin film was patterned into disk shaped features 500 nm in diameter on the top of the waveguide, and they were encapsulated with a 200 nm thick Al_2O_3 layer (see, inset of FIG. 13).

[0130] [5] The device made of the nanocomposite $Ge_4Sb_6Te_7$ thin film was found to be stable up to at least 30,000 cycles indicating the high reversibility of the crystallization and quenching processes of the coherent nanocomposite, (see, FIG. 4, panel b). The one-to-one comparison between the devices fabricated with $Ge_2Sb_2Te_5$ and $Ge_4Sb_6Te_7$ films here (see, FIG. 4, panels d and e) shows that $Ge_4Sb_6Te_7$ device exhibits up to 50% more in the number of non-volatile interval states than $Ge_2Sb_2Te_5$ device. The non-volatile symmetric switching was also realized (see, FIG. 13). These are important for the photonic memory and neuromorphic devices.

REFERENCES

[0131] 1. National Science and Technology Council (US). "Materials Genome Initiative for Global Competitiveness." Executive Office of the President, National Science and Technology Council (2011).

[0132] 2. Toher, C. et al. Unavoidable disorder and entropy in multi-component systems. *Npj Comput. Mater.* 5, 1-3 (2019).

[0133] 3. Hattrick-Simpers, J. R. et al. Perspective: Composition-structure-property mapping in high-throughput experiments: turning data into knowledge. *APL Mater* 4, 053211 (2016).

[0134] 4. Shapiro, E. et al. "Toward 2020 Science" Microsoft, (2006).

[0135] 5. Licklider, J. C. R. Man-computer symbiosis. *IRE Tran. Hum. Facto. Elect.* 1, 4-11 (1960).

[0136] 6. Shneiderman, B. Human-centered artificial intelligence: Reliable, safe & trustworthy. *Int. J. Hum. Comput. Inter.* 36, 495-504 (2020).

[0137] 7. Settles, B. Active learning literature survey. *Univ. Wis. Madison.* 52, 11 (2010).

[0138] 8. Jaynes, E. T. Bayesian methods: general background. (1986).

[0139] 9. Xue, D. et al. Accelerated search for materials with targeted properties by adaptive design. *Nat. Commun.* 7, 11241 (2016).

[0140] 10. Yuan, R. et al. Accelerated discovery of large electrostrains in BaTiO3-based piezoelectrics using active learning. *Adv. Mater.* 30, 1702884 (2018).

[0141] 11. Balachandran, P. V. et al. Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning. *Nat. Commun.* 9, 1668 (2018).

[0142] 12. Meredig, B. et al. Can machine learning identify the next high-temperature superconductor? Examining extrapolation performance for materials discovery. *Mol. Syst. Des. Eng.* 3, 819-825 (2018).

[0143] 13. Zhou, Z. et al. Optimization of molecules via deep reinforcement learning. *Sci. Rep.* 9, 1-10 (2019).

[0144] 14. Korovina, K. et al. Chembo: Bayesian optimization of small organic molecules with synthesizable recommendations. International Conference on Artificial Intelligence and Statistics. PMLR, 3393-3403 (2020).

[0145] 15. Li, Z. et al. Robot-accelerated perovskite investigation and discovery. *Chem. Mater.* 32, 5650 (2020).

[0146] 16. Pendleton, I. M. et al. Experiment specification, capture and laboratory automation technology (ESCALATE): a software pipeline for automated chemical experimentation and data management. *MRS Commun.* 9, 846-859 (2019).

[0147] 17. Zhong, M. et al. Accelerated discovery of CO2 electrocatalysts using active machine learning. *Nature* 581, 178-183 (2020).

[0148] 18. Nikolaev, P. et al. Autonomy in materials research: a case study in carbon nanotube growth. *Npj Comput. Mater.* 2, 16031 (2016).

[0149] 19. Roch, L. M. et al. ChemOS: orchestrating autonomous experimentation. *Sci. Robot.* 3, eaat5559 (2018).

[0150] 20. Attia, P. M. et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* 578, 397-402 (2020).

[0151] 21. Noack, M. M. et al. A kriging-based approach to autonomous experimentation with applications to x-ray scattering. *Sci. Rep.* 9, 1-19 (2019).

[0152] 22. Burger, B. et al. A mobile robotic chemist. *Nature* 583, 237-241 (2020).

[0153] 23. MacLeod, B. P. et al. Self-driving laboratory for accelerated discovery of thin-film materials. *Sci. Adv.* 6, eaaz8867 (2020).

[0154] 24. Langner, S. et al. Beyond ternary OPV: high-throughput experimentation and self-driving laboratories optimize multicomponent systems. *Adv. Mater.* 32, 1907801 (2020).

[0155] 25. Reyes, K. G. et al. The machine learning revolution in materials? *MRS Bulletin* 44, 530-537 (2019).

[0156] 26. Stein, H. S. et al. Progress and prospects for accelerating materials science with automated and autonomous workflows. *Chem. Sci.* 10, 9640-9649 (2019).

[0157] 27. Tabor, D. P. et al. Accelerating the discovery of materials for clean energy in the era of smart automation. *Nat Rev. Mater.* 3, 5-20 (2018).

[0158] 28. Jensen, K. F. et al. Autonomous discovery in the chemical sciences part I: Progress (Angewandte Chemie International Edition, 2019).

[0159] 29. Geiger, A. C. et al. Autonomous science: big data tools for small data problems in chemistry. *Mach. Learn. Chem.* 17, 450 (2020).

[0160] 30. Kolobov, A. V. et al. *Springer Handbook of Electronic and Photonic Materials,* (eds Kasap, S. & Capper, P.) (Springer International Publishing, Cham, 2017).

[0161] 31. Burr, G. W. et al. Neuromorphic computing using non-volatile memory. *Adv. Phys. X* 2, 89-124 (2016).

[0162] 32. Li, X. et al. Fast and reliable storage using a 5 bit, nonvolatile photonic memory cell. *Optica.* 6, 1-6 (2019).

[0163] 33. Giannopoulos, I. et al. 8-bit precision in-memory multiplication with projected phase-change memory. *IEEE International Electron Devices Meeting* (*IEDM*), 27.7.1-27.7.4 (2018)

[0164] 34. Simpson, R. E. et al. Interfacial phase-change memory. *Nat. Nanotechnol.* 6, 501-505 (2011).

[0165] 35. Ding, K. et al. Phase-change heterostructure enables ultralow noise and drift for memory operation. *Science* 366, 210-215 (2019).

[0166] 36. Kusne, A. G., "Materials Genome Initiative & Artificial Intelligence at NIST", DOE AMO Workshop on Artificial Intelligence Applied to Materials Discovery and Design, Pittsburgh, Pa. (2017).

[0167] 37. Green, M. et al. High-throughput experimental materials collaboratory, NIST materials genome initiative. https://mgi.nist.gov/htemc (2020)

[0168] 38. Srinivas, N. et al. Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Trans. Inf Theory.* 58, 3250-3265 (2012).

[0169] 39. Long, C. et al. Rapid structural mapping of ternary metallic alloy systems using the combinatorial approach and cluster analysis. *Rev. Sci. Instrum.* 78, 072217 (2007).

[0170] 40. spec Software for Diffraction, Certified Scientific Software. https://certif.com/spec.html. (2017)

[0171] 41. Curtarolo, S. et al. AFLOWLIB.ORG: a distributed materials properties repository from high-throughput ab initio calculations. *Comput. Mater. Sci.* 58, 227-235 (2012).

[0172] 42. Kusne, A. G. et al. On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets. *Sci. Rep.* 4, 6367 (2014).

[0173] 43. Ren, F. et al. On-the-fly data assessment for high-throughput X-ray diffraction measurements. *ACS Combi. Sci.* 19, 377-385 (2017).

[0174] 44. Kusne, A. G. et al. High-throughput determination of structural phase diagram and constituent phases using GRENDEL. *Nanotechnology.* 26, 444002 (2015).

[0175] 45. Dobruschin, P. L. The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory Probab. Its Appl* 13, 197-224 (1968).

[0176] 46. Zhu, X. et al. Semi-supervised learning using gaussian fields and harmonic functions. *Proceedings of the* 20*th International conference on Machine learning* (*ICML*-03) (2003).

[0177] 47. Brochu, E. et al. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv1012.2599. Preprint at https://arxiv.org/abs/1012.2599 (2010).

[0178] 48. Hastie, T. et al. The Elements of Statistical Learning 2nd Ed. (Springer, 2009).

[0179] 49. Shportko, K. et al. Resonant bonding in crystalline phase-change materials. *Nat. Mater.* 7, 653-658 (2008).

[0180] 50. Rao, F. et al. Reducing the stochasticity of crystal nucleation to enable subnanosecond memory writing. *Science.* 358, 1423-1427 (2017).

[0181] 51. Zhu, M. et al. One order of magnitude faster phase change at reduced power in Ti—Sb—Te. *Nat. Commun.* 5, 1-6 (2014).

[0182] 52. Deringer, V. L. et al. Bonding nature of local structural motifs in amorphous GeTe. Angew. *Chem. Int. Ed* 53, 10817-10820 (2014).

[0183] 53. Wuttig, M. et al. Phase-change materials for rewriteable data storage. *Nat. Mater.* 6, 824-832 (2007).

[0184] 54. Lencer, D. et al. A map for phase-change materials. *Nat. Mater.* 7, 972-977 (2008).

[0185] 55. Wuttig, M. et al. Phase-change materials for non-volatile photonic applications. *Nat. Photonics.* 11, 465 (2017).

[0186] 56. Pries, J. et al. Phase-change materials: Empowered by an unconventional bonding mechanism. *MRS Bull* 44, 699-704 (2019).

[0187] The foregoing is illustrative of the present disclosure and is not to be construed as limiting thereof. Further embodiments of the present disclosure are exemplified in the following claims.

[0188] All patents, patent applications, and printed publications cited herein are incorporated herein by reference in the entireties, except for any definitions, subject matter disclaimers or disavowals, and except to the extent that the incorporated material is inconsistent with the express disclosure herein, in which case the language in this disclosure controls.

1. A phase-change memory material comprising a material of formula (I):

$$TM_xSb_yTe_z \qquad (I)$$

wherein TM is a transition metal, wherein x, y, and z are each independently an integer between 1-10, and wherein the phase-change memory material has an extinction difference between crystalline and amorphous phases ($k_c - k_a$) greater than about 2 in a wavelength range of about 1000 nm to about 1500 nm.

2. The phase-change memory material of claim 1, wherein TM of formula (I) is Ge.

3. The phase-change memory material of claim 2, wherein the material of formula (I) is $Ge_4Sb_6Te_7$.

4. The phase-change memory material of claim 1, wherein the phase-change memory material has a ΔE_g of greater than about 0.76 eV.

5-7. (canceled)

8. A method of forming a photonic switching device comprising forming a film on a substrate, the film comprising the phase-change memory material of claim 1.
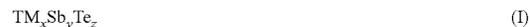
9-14. (canceled)

15. A method of forming a memory device comprising forming a film on a substrate, the film comprising a phase-change memory material selected to have an extinction difference between crystalline and amorphous phases ($k_c - k_a$) greater than about 2 in a wavelength range of about 1000 nm to about 1500 nm.

16. A method of forming a memory device comprising:
selecting a phase-change memory material having an extinction difference between crystalline and amorphous phases ($k_c - k_a$) greater than about 2 in a wavelength range of about 1000 nm to about 1500 nm;
forming a lower electrode on a substrate;
forming a film including the phase-change memory material on a lower electrode; and
forming an upper electrode on the film.

17. The method of claim 16, wherein selecting the phase-change memory material comprises evaluating the extinction difference of the phase-change memory material via Bayesian active learning.

18. The method of claim 17, wherein the Bayesian active learning comprises closed-loop autonomous materials exploration and optimization (CAMEO).

19. The method of claim 16, wherein the phase-change memory material comprises a material of formula (I):

$$TM_xSb_yTe_z \qquad (I)$$

wherein TM is a transition metal, and wherein x, y, and z are each independently an integer between 1-10.

20. The method of claim 19, wherein TM is Ge.

21. The method of claim 19, wherein the material of formula (I) comprises $Ge_4Sb_6Te_7$.

22. (canceled)

23. A method of selecting a phase-change memory material comprising:
evaluating optical bandgap difference of phase-change memory materials; and
selecting a phase-change memory material or materials having larger or largest optical bandgap differences,
wherein evaluation is conducted via Bayesian active learning.

24. The method of claim 23, wherein the Bayesian active learning comprises closed-loop autonomous materials exploration and optimization (CAMEO).

25. The method of claim 23, wherein the phase-change memory materials comprise materials of formula (I):

$$TM_xSb_yTe_z \qquad (I)$$

wherein TM is a transition metal, and wherein x, y, and z are independently integers between 1-10.

26. The method of claim 25, wherein TM is Ge.

27. The method of claim 25, wherein the materials of formula (I) comprise $Ge_4Sb_6Te_7$.

28. The method of claim 25, wherein the phase-change memory materials have a ΔE_g of greater than about 0.76 eV.

29-30. (canceled)

31. The phase-change memory material of claim 1, comprising a material of formula (II):

$$Ge_xSb_yTe_z \qquad (I).$$

32. The phase-change memory material of claim 31, wherein x is between about 3.4 and 4, y is between about 5 and 6, and z is between about 6.5 and 7.5.

* * * * *