

## MIT Open Access Articles

*Origins of hole traps in hydrogenated nanocrystalline and amorphous silicon revealed through machine learning*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Mueller, Tim, Eric Johlin, and Jeffrey C. Grossman. "Origins of Hole Traps in Hydrogenated Nanocrystalline and Amorphous Silicon Revealed through Machine Learning." Phys. Rev. B 89, no. 11 (March 2014). © 2014 American Physical Society

**As Published:** <http://dx.doi.org/10.1103/PhysRevB.89.115202>

**Publisher:** American Physical Society

**Persistent URL:** <http://hdl.handle.net/1721.1/88769>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of Use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



# Origins of hole traps in hydrogenated nanocrystalline and amorphous silicon revealed through machine learning

Tim Mueller,<sup>1</sup> Eric Johlin,<sup>2</sup> and Jeffrey C. Grossman<sup>2,\*</sup>

<sup>1</sup>*Department of Materials Science and Engineering, Johns Hopkins University, Baltimore Maryland 21218, USA*

<sup>2</sup>*Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge Massachusetts 02139, USA*

(Received 29 October 2013; published 10 March 2014)

Genetic programming is used to identify the structural features most strongly associated with hole traps in hydrogenated nanocrystalline silicon with very low crystalline volume fraction. The genetic programming algorithm reveals that hole traps are most strongly associated with local structures within the amorphous region in which a single hydrogen atom is bound to two silicon atoms (bridge bonds), near fivefold coordinated silicon (floating bonds), or where there is a particularly dense cluster of many silicon atoms. Based on these results, we propose a mechanism by which deep hole traps associated with bridge bonds may contribute to the Staebler-Wronski effect.

DOI: [10.1103/PhysRevB.89.115202](https://doi.org/10.1103/PhysRevB.89.115202)

PACS number(s): 81.05.Gc, 31.15.A–, 07.05.Mh

## I. INTRODUCTION

Hydrogenated amorphous silicon (*a*-Si:H) is a leading material for thin-film photovoltaics due to its low cost, low toxicity, and greater optical absorption relative to crystalline silicon [1–3]. However, its efficiency is hampered by low hole mobility and the Staebler-Wronski effect [3,4], in which light-induced defects trap carriers and lead to performance degradation over time. These problems can be mitigated through the use of nanocrystalline (also known as microcrystalline) silicon (*nc*-Si:H), which consists of small silicon crystals connected by regions of hydrogenated amorphous silicon [5–9]. There have been several theoretical studies on the nature of hole traps in *a*-Si:H [10–17], but despite the technological importance of *nc*-Si:H the nature of hole traps in this material remains relatively poorly understood. We have used density functional theory (DFT) [18,19] to identify the deepest hole traps in an ensemble of 1045 *nc*-Si:H structures, and we have analyzed these results using genetic programming [20] to determine which structural features contribute most significantly to the depth of hole traps. Our results reveal that holes are less likely to be trapped near the crystalline region than in the amorphous region. The genetic programming algorithm has identified three structural features—each of which is found in the amorphous region of the material—that most strongly affect the depth of hole traps. One of these features, Si-H-Si bridge bonds, has been predicted to become more prevalent during light soaking [21–23] and hence we propose that it may contribute to the Staebler-Wronski effect.

## II. METHODS

### A. Sample generation

The label *nc*-Si:H is broadly applied to materials that range from nearly amorphous to polycrystalline. The degree of crystallinity in the material can be tailored by adjusting the amount of hydrogen dilution in the silane mixture used during vapor deposition. As the concentration of hydrogen is

increased, the material goes from being purely amorphous, to a mix of crystalline and amorphous phases, to a polycrystalline phase [24,25]. Experimental studies suggest that the material with the best properties for use in solar cells is near the transition from purely amorphous to mixed phases [26–28]. Such a material is nearly amorphous, but with a higher degree of structural order than a purely amorphous material. We have modeled materials with a very small crystalline volume about 0.6 nm wide, with crystalline volume fraction of about 4%, and about 1.1 nm of amorphous material separating the crystalline regions (Fig. 1). These materials enable us to explore, with reasonable computational cost, how an incremental increase in local crystalline order affects hole traps in hydrogenated amorphous silicon.

To explore the variety of possible local environments in the amorphous region of the material, we have generated a library of 1045 different structures, in which each unit cell contains 216 silicon atoms and 20 hydrogen atoms. This concentration of hydrogen is identical to the concentration used in our study of amorphous silicon [10], allowing for direct comparison with those results. The amorphous structures were generated using the same validated Wooten-Winer-Weaire (WWW) method [29] we previously used to generate the amorphous silicon structures [10]. A Keating potential [30] was used to determine the energy of the bond changes. Hydrogen was introduced to the cell prior to annealing by replacing randomly selected Si-Si bonds with pairs of Si-H bonds. The crystalline region was created by freezing the positions of the ten silicon atoms within 3 Å of the center of the cubic unit cell during the WWW process. The resulting structures were analyzed to ensure that they consisted of a small, 0.6-nm crystalline core surrounded by hydrogenated amorphous silicon.

### B. Hole trap depth calculation

Hole trap depths were calculated using density functional theory (DFT) [18,19] as implemented in the SIESTA software package [31] using a double-zeta polarized (DZP) basis and the Perdew-Burke-Ernzerhof (PBE) [32] exchange correlation functional. We have used the same approach described in Ref. [10]. The hole trap depths were calculated by removing a single electron per unit cell and calculating the energy

\*Author to whom correspondence should be addressed: [jcg@mit.edu](mailto:jcg@mit.edu)

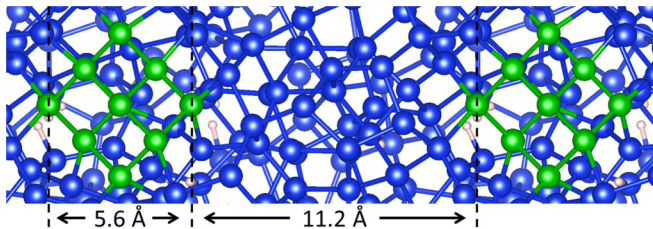


FIG. 1. (Color online) A sample *nc*-Si:H structure. Blue spheres represent amorphous Si, green represent crystalline Si, and white represent H.

difference between the charged sample and a neutral sample. This approach to calculating hole trap depths accounts for self-trapping due to structural relaxation around the holes, which increased the average hole trap depth in our samples by 142 meV. To ensure that we were calculating the depths of reversible hole traps, charge and discharge cycles were repeated until the calculated trap depth converged. Spin-polarized calculations were used for all charged samples. Non-spin-polarized calculations were run on the neutral samples, and we confirmed that this did not significantly affect the results by recalculating the final neutral structures with spin polarization allowed. All atomic positions, lattice parameters, and cell shapes were fully relaxed. The Brillouin zone was sampled at the  $\Gamma$  point.

The locations of the trapped holes were determined by calculating the difference between the electronic charge densities of the material in the neutral and charged states. This approach captures both the positive shift in local charge due to the presence of a hole and the effects of localized displacements in atomic positions. We have previously showed that such displacements are commonly observed for hole traps, and are particularly large for deep hole traps [10]. One detail to be considered when comparing the charge density in the neutral and charged states is that there is usually a slight difference in the Bravais lattices for these structures. We compensated for the lattice change by applying a linear transformation to each charged structure to align it with the corresponding neutral structure. The difference in the lattice between the neutral and charged structures was typically small; the root-mean-square difference between the eigenvalues for these linear transformations and the eigenvalues of the identity matrix is less than 0.001.

### C. Data analysis (genetic programming)

It can be difficult to manually determine how local structure affects the depth of a hole trap. Previous computational research on *a*-Si:H has examined the presence of undercoordinated Si (dangling bonds) [10–12,15,17], overcoordinated Si (floating bonds) [10], and strained silicon bonds [11,14,16,17]. It has become increasingly clear that a variety of local structural environments can trap holes, and not all of these are well understood. We have used a different approach to identify the structural features that are most strongly associated with hole traps. Rather than investigate a short list of simple, intuitive structural features, we have calculated the values of 242 structural descriptors for each sample in

our training set. These descriptors include simple values such as unit cell volume to more complex values such as the third-largest distance in the cell between a silicon atom and its fourth-nearest silicon neighbor. A complete description of the descriptors is provided in the Supplemental Material [33]. When developing the list of descriptors, we have been mindful of the fact that because the hole is trapped at the most energetically favorable location in the material, it is important to consider extreme values of local structural variables rather than the average values.

To determine which descriptors are most strongly associated with hole traps, we have used genetic programming as implemented in the EUREQA software package [34]. Genetic programming is a machine learning method that has been successfully used in a variety of fields [20], and here we demonstrate how it can be used to identify atomic-scale descriptors for materials. In our genetic programming approach, an evolutionary algorithm is used to identify the functions that best predict hole trap depth. Using the set of 242 descriptors as variables, a population of functions is created by combining basic mathematical operators and functions (e.g., addition, subtraction, minimum value). The fitness of these functions is evaluated by calculating the correlation between the function output values and the DFT-calculated hole trap depths. The least fit functions are discarded, and from the remaining functions a new population of functions is generated using both crossover operations, in which elements of two existing functions are mixed to create new functions, and mutation operations, in which an element of a function is randomly changed (Fig. 2). The cycle is repeated for many generations, resulting in a population of functions that are able to predict trends in hole trap depths from atomic-level structural information. In our genetic programming calculations, no preprocessing or renormalization was done for any of the descriptors used as input variables.

It is necessary to take care not to overfit the training data when using genetic programming. Simple functions with few degrees of freedom are less likely than complex functions to overfit the data. However, accurate prediction of hole trap depths may require the use of complex functions. To balance this tradeoff, complexity values were assigned to each “building block” used to construct the functions, where

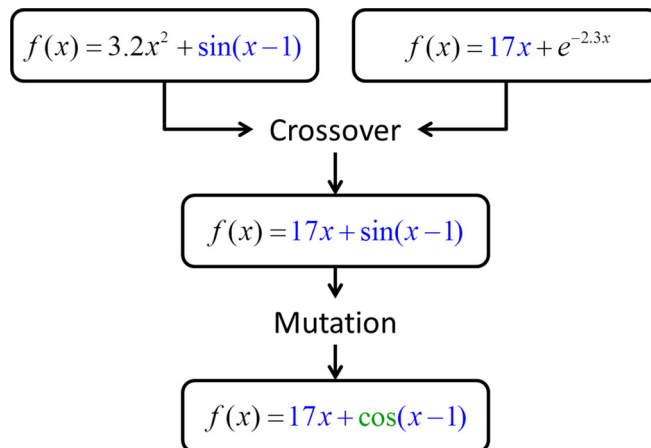


FIG. 2. (Color online) Examples of possible crossover and mutation steps in genetic programming.

TABLE I. Allowed building blocks in the genetic programming algorithm.

Building block	Complexity value
Constant value	0
Input variable	1
Addition	1
Subtraction	1
Multiplication	1
Division	2
Negation	1
Exponential	4
Natural logarithm	4
Power	5
Square root	4
Logistic function	2
Minimum	0
Maximum	0
Absolute value	2

the building blocks are simple functions or operators. The complexity of each generated function was calculated by adding up the complexity values of the building blocks used to construct the function. The list of allowed building blocks and complexity values is provided in Table I. The “min” and “max” functions were assigned complexity values of zero, due to the likelihood that these building blocks would be instrumental in identifying the deepest hole trap in each sample. The best functions with respect to fitness and complexity were identified as those that existed on the Pareto frontier [35], defined as the set of functions for which no other function is both more fit and less complex. The default algorithm in EUREQA was used to split the samples randomly into a training set (used to search for new functions), and a validation set (used to determine the Pareto frontier).

### III. RESULTS

Sixteen functions were identified on the Pareto frontier. The simplest function, with a correlation coefficient of 0.38, is the minimum of 1.62501 Å and the longest distance between H and its nearest-neighbor Si. 1.62501 Å is approximately the length that distinguishes bridge-bonded hydrogen from a Si-H/dangling-bond pair [36–38], indicating that deep hole traps are associated with bridge bonds. The next simplest function, with a correlation coefficient of 0.57, is proportional to a measure of the maximum coordination of a silicon atom, indicating a floating bond. The maximum correlation coefficient with the DFT-calculated trap depths was 0.65, obtained using the following function (Fig. 3):

$$\frac{\min(1.66355, a) \max(5.37551, c) - f - bd}{g} - h \max(3.42929, e), \quad (1)$$

where the variables are defined in Table II.

The equations discovered by the genetic programming algorithm provide insight into what structural features are most strongly associated with hole traps. The most relevant descriptors, ranked by the number of times each variable

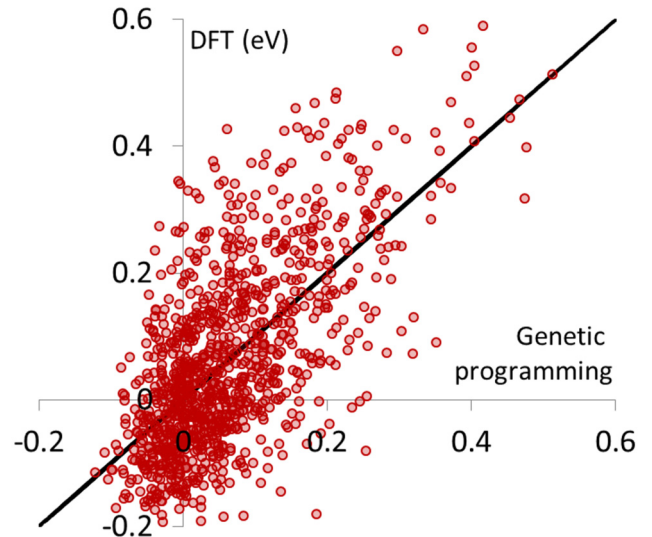


FIG. 3. (Color online) The DFT-calculated trap depths compared to those predicted by the function discovered by genetic programming. The black line indicates an ideal match.

appears on the Pareto frontier, are shown in Table II. These descriptors fall into three general categories. The first set of descriptors is associated with Si-H-Si bridge bonds, the second set indicates overcoordinated silicon atoms (“floating bonds”), and the third represents regions of densely packed silicon. These features exist in the amorphous regions of the material, providing additional evidence that holes are more likely to be trapped in this region. To confirm this result we ran the genetic programming algorithm on our previously generated [10] library of 2700 *a*-Si:H structures. The most common descriptors on the Pareto frontier fall into the same three general categories as for the *nc*-Si:H samples (Table III), providing further evidence that these three categories are particularly relevant to hole trap depths in amorphous regions. The two variables in bold in Table III are identical to variables found on the Pareto frontier of the *nc*-Si:H samples, and many of the others are similar to variables found for the *nc*-Si:H samples. Descriptors that indicate the presence of isolated dangling bonds show up on neither list, suggesting that dangling bonds are not one of the features most strongly associated with the deepest hole trap in each sample. This result supports our previous results on *a*-Si:H, in which the presence of dangling bonds was found to have little effect on the distribution of hole trap depths [10].

The identification of floating bonds as significant contributors to hole trap depths is consistent with our previous analysis of *a*-Si:H [10]. Floating bonds, indicated by such descriptors as the shortest distance to the fifth-nearest neighbor silicon atom, are commonly found in structures with regions of dense silicon, which are indicated by descriptors such as the fourth-shortest distance between a silicon atom and its eighth-nearest neighbor silicon. The importance of such seemingly obscure descriptors is consistently identified by the genetic programming algorithm. The identification of regions of dense silicon as a predictor of hole trap depths is consistent with previous studies identifying valence band-edge states with local regions of short Si-Si bonds [12–14,16].

TABLE II. The variables in the functions on the Pareto frontier for the *nc*-Si:H samples.

Variable	Count <sup>a</sup>	Symbol <sup>b</sup>	BB <sup>c</sup>	FB <sup>d</sup>	DS <sup>e</sup>
The largest distance between a H atom and its nearest Si neighbor	13	<i>a</i>	X		
The shortest distance between a Si atom and its sixth-nearest Si neighbor	12	<i>b</i>			X
The maximum bond valence sum on a Si atom	11	<i>c</i>		X	
The smallest value for the fifth-smallest relative bond length around a Si atom	10	<i>d</i>		X	
The fourth-shortest distance between a Si atom and its eighth-nearest neighbor	8	<i>e</i>		X	
The shortest distance between a Si atom and its fifth-nearest neighbor	6			X	
The second-shortest distance between a Si atom and its fifth-nearest neighbor	6	<i>f</i>		X	X
The third-shortest distance between a Si atom and its sixth-nearest neighbor	3	<i>g</i>			X
The H-Si nearest-neighbor distance for the hydrogen atom with the fourth-smallest difference between the distances to the two Si atoms nearest to a H atom	2	<i>h</i>	X		
The fifth-shortest distance between a Si atom and its fifth-nearest neighbor	1			X	X

<sup>a</sup>Count is the number of times the variable appears in a function on the Pareto frontier.

<sup>b</sup>Symbol is the symbol for the variable used in Eq. (1).

<sup>c</sup>BB indicates that the hole trap is associated with a bridge bond.

<sup>d</sup>FB indicates that the hole trap is associated with a floating bond.

<sup>e</sup>DS indicates that the hole trap is associated with a region of dense silicon.

The relevance of Si-H-Si bridge bonds is of particular interest. Of the three categories of structural features associated with hole trap depths, it is the only one that directly involves hydrogen, and we had not considered bridge bonds to be a likely feature of hole traps prior to the machine learning results. The effect of bridge bonds on the distribution of hole trap depths is shown in Fig. 4. For this plot, bridge bonds were identified using a descriptor discovered by the genetic programming algorithm: a hydrogen atom in which the nearest silicon atom is more than 1.625 Å away. The peak of the distribution of structures containing bridge bonds is about 0.3 eV deeper than for the distribution of all structures. The average hole trap in structures containing bridge bonds is about 0.2 eV deeper than that of structures containing no bridge bonds.

Although numerous descriptors are associated with bridge bonds, the variable representing the longest distance in the cell between a hydrogen atom and its nearest-neighbor silicon atom is particularly relevant to predicting hole trap depth. The structures with the eight deepest hole traps in our

1045-structure data set all have unusually high values for this variable, within the top 55 structures. To determine the location of the trapped holes in these structures, we calculated the difference between the total charge densities of the materials in the neutral and charged states. These comparisons confirm that the holes in these structures are trapped around the bridge bond. Two representative charge density plots can be found in Fig. 5. A common feature for these deep hole traps is for the hole to be spread over other nearby defects, such as overcoordinated silicon. This observation, as well as functions such as Eq. (1), suggest that interactions between multiple defects lead to particularly deep traps. We note that the extent of the delocalization of the hole might be exaggerated due to the self-interaction error in the generalized gradient approximation [39].

#### IV. DISCUSSION

Deep hole traps associated with bridge bonds may contribute to the Staebler-Wronski effect. It has been proposed

TABLE III. The variables in the functions on the Pareto frontier for the *a*-Si:H samples.

Variable	Count <sup>a</sup>	BB <sup>b</sup>	FB <sup>c</sup>	DS <sup>d</sup>
<b>The shortest distance between a silicon atom and the fifth-nearest silicon atom</b>	<b>15</b>		X	
The fourth-shortest distance between a silicon atom and the fifth-nearest silicon atom	11		X	X
<b>The longest distance between a hydrogen atom and the nearest silicon atom</b>	<b>10</b>	X		
The second-shortest distance between a silicon atom and the eighth-nearest silicon atom	8			X
The second-shortest distance between a silicon atom and the sixth-nearest silicon atom	8			X
The second-highest valence sum for any silicon atom	7		X	X
The second-smallest difference between the distances from a H atom to its first- and second-nearest neighbor Si atoms	6	X		
The root-mean-square deviation between the bond valence sums for hydrogen atoms and 1	3	X		
The smallest value for the shortest relative bond length around a silicon atom	1		X	
The fourth-longest distance between a silicon atom and the seventh-nearest silicon atom	1			X
The root-mean-square difference between the bond valence sums for silicon atoms and 4	1		X	X

<sup>a</sup>Count is the number of times the variable appears in a function on the Pareto frontier.

<sup>b</sup>BB indicates that the hole trap is associated with a bridge bond.

<sup>c</sup>FB indicates that the hole trap is associated with a floating bond.

<sup>d</sup>DS indicates that the hole trap is associated with a region of dense silicon.

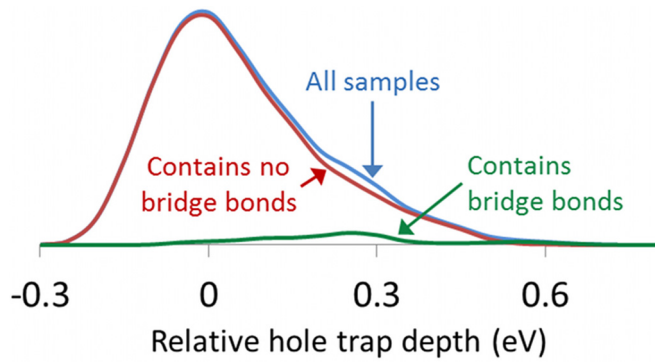


FIG. 4. (Color online) The distribution of hole trap depths for all samples, only samples with no bridge bonds and only samples containing bridge bonds. The trap depths are calculated relative to the peak of the distribution for all samples, with more positive numbers indicating deeper traps.

that the creation of dangling bonds upon light soaking may be accompanied by the creation of bridge bonds through a reaction in which hydrogen is inserted into a weakened Si-Si bond, leaving behind a dangling bond [40]. The hydrogen is then expected to become relatively mobile, moving through a series of metastable Si-H-Si configurations [38,41–43], consistent with the experimentally observed increase in the diffusivity of H upon exposure to light [44]. Kinetic models

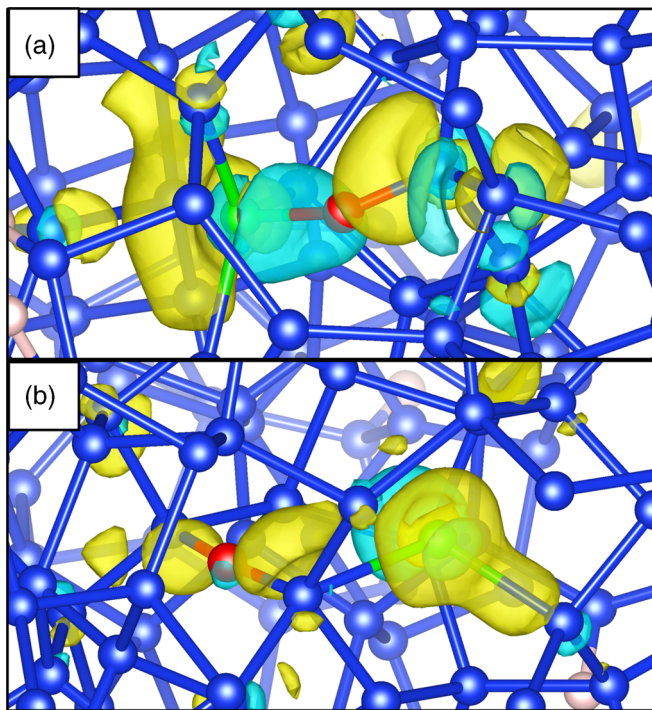


FIG. 5. (Color online) The (a) first- and (b) fifth-deepest hole traps in the sample. Blue spheres represent silicon, white spheres represent hydrogen, and red spheres represent bridge-bonded hydrogen. Yellow contours mark regions in which the electron charge density decreases in the presence of a hole, and blue contours mark regions in which it increases. The green spheres represent (a) a silicon atom with only three silicon neighbors, and (b) a silicon atom that has two silicon neighbors within 2.5 Å and eight silicon neighbors within 3 Å.

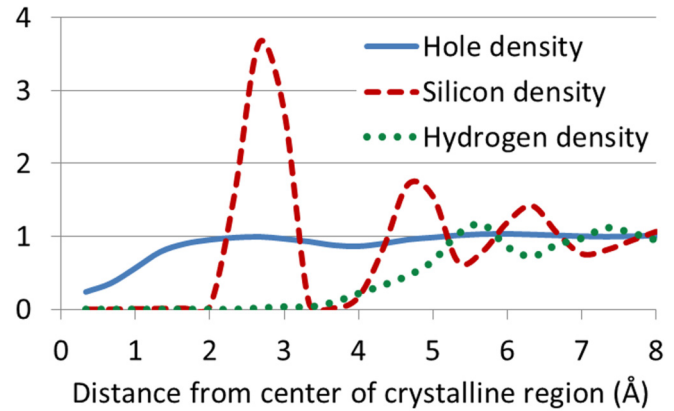


FIG. 6. (Color online) The normalized distributions of holes, silicon atoms, and hydrogen atoms measured from the center of the crystalline region. A value of 1 represents the average value for the material.

predict that the concentration of Si-H-Si configurations should increase rapidly upon light soaking [21–23]. If bridge bonds are associated with deep hole traps as we predict, an increase in the bridge-bond concentration due to light soaking would likely result in the degradation of device performance.

A similar theory has been put forth by Morigaki [45], in which the Si-H-Si structures are described as hydrogen-related dangling bonds in which the hydrogen atom is preferentially bound to one silicon atom. This model has been challenged by electron-spin-resonance studies indicating that there are no hydrogen atoms within 3 Å of dangling bonds in *a*-Si:H [46,47], but more recent studies have called this conclusion into question [48–50]. About 2% of the structures in our sample, and three of the eight deepest hole traps, contain bridge bonds in which the difference between the two H-Si bond lengths is less than 0.05 Å. This suggests that the hydrogen is nearly equally shared between the two silicon atoms, and it may be more appropriate to think of such configurations as overcoordinated hydrogen than as hydrogen-related dangling bonds.

Biswas *et al.* have previously proposed that a positively charged bond-centered hydrogen might contribute to the Staebler-Wronski effect by converting to Si-H and a dangling bond upon electron capture during light soaking [23,36]. Some evidence for the reverse process can be seen in the structure shown in Fig. 5(a) in which the Si-H bond lengths in the positively charged sample (1.69 and 1.77 Å) are more evenly matched than the bond lengths in the neutral sample (1.67 and 1.84 Å). However, in the structures with the eight deepest hole traps in our sample, we see little change (~0.01 Å) in the average H-Si bond lengths in bridge bonds upon switching between neutral and positively charged states, indicating that any significant bonding rearrangement would be thermally activated.

The results from the genetic algorithm suggest that holes are unlikely to be trapped near the crystalline region of the material. This is confirmed by plotting the normalized average absolute value of the charge difference as a function of the distance to the center of the crystalline region (Fig. 6). The hole density is lowest in the center of the crystalline volume and greatest in the regions farthest from the crystalline region. The slight dip in density observed near the transition

from the crystalline to amorphous region is likely related to the dip in the radial distribution function for silicon. However, at no point in or near the crystalline region does the hole density fluctuate above the average value for the material. A comparison with our previous results on purely amorphous silicon [10] reveals that the average trap depth is 16 meV lower in the *nc*-Si:H samples, a difference that is statistically significant with 99.7% certainty.

## V. CONCLUSION

The complexity of *a*-Si:H and *nc*-Si:H is made apparent by the fact that the best functions found by the genetic programming algorithm were only able to account for about half of the variance in the hole trap depths. It is appealing to ascribe the electronic states in amorphous materials to simple features such as dangling bonds, floating bonds, and bridge bonds, but such simple features were not found to be universally predictive of hole trap depths. This is perhaps not surprising considering theoretical studies, including this one, which show that holes are often trapped in states that

extend over many atoms [10–14,17]. Despite this complexity, the genetic programming algorithm was able to successfully identify two structural features that previous computational studies had associated with hole traps, as well as one additional feature. The descriptors identified by the genetic programming algorithm have enabled us to identify a possible mechanism for the widely studied Staebler-Wronski effect. These results demonstrate that genetic programming is a powerful tool for the identification of structural descriptors in materials and can facilitate a qualitative understanding of complex physical processes.

## ACKNOWLEDGMENTS

This research was sponsored by King Fahd University of Petroleum and Minerals in Dhahran, Saudi Arabia, through the Center for Clean Water and Clean Energy at MIT and KFUPM under Project No. R1-CE-08, and by the National Science Foundation under Grant No. 1035400. Structure images were generated using VESTA [51]. Supercomputing resources provided by the National Energy Research Scientific Computing Center were used in this research.

- 
- [1] M. Konagai, *Jpn. J. Appl. Phys.* **50**, 030001 (2011).
  - [2] L. V. Mercaldo, M. L. Addonizio, M. D. Noce, P. D. Veneri, A. Scognamiglio, and C. Privato, *Appl. Energy* **86**, 1836 (2009).
  - [3] V. Avrutin, N. Izyumskaya, and H. Morkoç, *Superlattices Microstruct.* **49**, 337 (2011).
  - [4] D. L. Staebler and C. R. Wronski, *Appl. Phys. Lett.* **31**, 292 (1977).
  - [5] T. Dylla, F. Finger, and E. A. Schiff, *Appl. Phys. Lett.* **87**, 032103 (2005).
  - [6] I.-C. Cheng and S. Wagner, *Appl. Phys. Lett.* **80**, 440 (2002).
  - [7] T. Dylla, S. Reynolds, R. Carius, and F. Finger, *J. Non-Cryst. Solids* **352**, 1093 (2006).
  - [8] H.-n. Liu and M.-d. Xu, *Solid State Commun.* **58**, 601 (1986).
  - [9] A. V. Shah, J. Meier, E. Vallat-Sauvain, N. Wyrsh, U. Kroll, C. Droz, and U. Graf, *Sol. Energy Mater. Sol. Cells* **78**, 469 (2003).
  - [10] E. Johlin, L. K. Wagner, T. Buonassisi, and J. C. Grossman, *Phys. Rev. Lett.* **110**, 146805 (2013).
  - [11] L. K. Wagner and J. C. Grossman, *Phys. Rev. Lett.* **101**, 265501 (2008).
  - [12] Y. Pan, F. Inam, M. Zhang, and D. A. Drabold, *Phys. Rev. Lett.* **100**, 206403 (2008).
  - [13] D. A. Drabold, Y. Li, B. Cai, and M. Zhang, *Phys. Rev. B* **83**, 045201 (2011).
  - [14] P. A. Fedders, D. A. Drabold, and S. Nakhmanson, *Phys. Rev. B* **58**, 15624 (1998).
  - [15] K. Jarolimek, R. A. de Groot, G. A. de Wijs, and M. Zeman, *Phys. Rev. B* **79**, 155206 (2009).
  - [16] Y. Pan, M. Zhang, and D. A. Drabold, *J. Non-Cryst. Solids* **354**, 3480 (2008).
  - [17] P. A. Khomyakov, W. Andreoni, N. D. Afify, and A. Curioni, *Phys. Rev. Lett.* **107**, 255502 (2011).
  - [18] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
  - [19] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
  - [20] R. Poli, W. B. Langdon, and N. F. McPhee, *A Field Guide to Genetic Programming* (Lulu.com, 2008), with contributions by J. R. Koza. Available at <http://www.gp-field-guide.org.uk>.
  - [21] K. Morigaki and H. Hikita, *J. Non-Cryst. Solids* **266–269**, 410 (2000).
  - [22] K. Morigaki and H. Hikita, *Solid State Commun.* **114**, 69 (2000).
  - [23] H. M. Branz, *Phys. Rev. B* **59**, 5498 (1999).
  - [24] D. V. Tsu, B. S. Chao, S. R. Ovshinsky, S. Guha, and J. Yang, *Appl. Phys. Lett.* **71**, 1317 (1997).
  - [25] R. W. Collins, A. S. Ferlauto, G. M. Ferreira, C. Chen, J. Koh, R. J. Koval, Y. Lee, J. M. Pearce, and C. R. Wronski, *Sol Energy Mater. Sol. Cells* **78**, 143 (2003).
  - [26] R. W. Collins, J. Koh, A. S. Ferlauto, P. I. Rovira, Y. Lee, R. J. Koval, and C. R. Wronski, *Thin Solid Films* **364**, 129 (2000).
  - [27] J. M. Pearce, N. Podraza, R. W. Collins, M. M. Al-Jassim, K. M. Jones, J. Deng, and C. R. Wronski, *J. Appl. Phys.* **101**, 114301 (2007).
  - [28] O. Vetterl, F. Finger, R. Carius, P. Hapke, L. Houben, O. Kluth, A. Lambertz, A. Mück, B. Rech, and H. Wagner, *Sol Energy Mater. Sol. Cells* **62**, 97 (2000).
  - [29] F. Wooten, K. Winer, and D. Weaire, *Phys. Rev. Lett.* **54**, 1392 (1985).
  - [30] P. N. Keating, *Phys. Rev.* **145**, 637 (1966).
  - [31] M. S. José, A. Emilio, D. G. Julian, G. Alberto, J. Javier, O. Pablo, and S.-P. Daniel, *J. Phys.: Condens. Matter* **14**, 2745 (2002).
  - [32] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
  - [33] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.89.115202> for the complete list of descriptors.

- [34] M. Schmidt and H. Lipson, *Science* **324**, 81 (2009).
- [35] G. F. Smits and M. Kotanchek, in *Genetic Programming Theory and Practice II*, edited by U.-M. O'Reilly, T. Yu, R. Riolo, and B. Worzel (Springer, New York, 2005), p. 283.
- [36] R. Biswas, I. Kwon, and C. M. Soukoulis, *Phys. Rev. B* **44**, 3403 (1991).
- [37] B. Tuttle and J. B. Adams, *Phys. Rev. B* **57**, 12859 (1998).
- [38] S. Sriraman, S. Agarwal, E. S. Aydil, and D. Maroudas, *Nature* **418**, 62 (2002).
- [39] G. Pacchioni, F. Frigoli, D. Ricci, and J. A. Weil, *Phys. Rev. B* **63**, 054102 (2000).
- [40] M. Stutzmann, W. B. Jackson, and C. C. Tsai, *Phys. Rev. B* **32**, 23 (1985).
- [41] Y. S. Su and S. T. Pantelides, *Phys. Rev. Lett.* **88**, 165503 (2002).
- [42] S. Lanzavecchia and L. Colombo, *Europhys. Lett.* **36**, 295 (1996).
- [43] T. A. Abtew and D. A. Drabold, *Phys. Rev. B* **74**, 085201 (2006).
- [44] P. V. Santos, N. M. Johnson, and R. A. Street, *Phys. Rev. Lett.* **67**, 2686 (1991).
- [45] K. Morigaki, *Jpn. J. Appl. Phys.* **27**, 163 (1988).
- [46] T. Shimizu, *Jpn. J. Appl. Phys.* **43**, 3257 (2004).
- [47] S. Yamasaki, T. Umeda, J. Isoya, J. H. Zhou, and K. Tanaka, *J. Non-Cryst. Solids* **227–230**, 332 (1998).
- [48] M. Fehr, A. Schnegg, C. Teutloff, R. Bittl, O. Astakhov, F. Finger, B. Rech, and K. Lips, *Phys. Status Solidi A* **207**, 552 (2010).
- [49] O. Astakhov, R. Carius, P. Yu, V. Borysenko, D. Barankov, and F. Finger, *J. Phys.: Condens. Matter* **24**, 305801 (2012).
- [50] G. Pfanner, C. Freysoldt, J. Neugebauer, F. Inam, D. Drabold, K. Jarolimek, and M. Zeman, *Phys. Rev. B* **87**, 125308 (2013).
- [51] K. Momma and F. Izumi, *J. Appl. Crystallogr.* **41**, 653 (2008).