On the Prediction of Ternary Semiconductor Properties by Artificial Intelligence Methods

Yingzhi Zeng,† Soo Jin Chua,‡ and Ping Wu*,†

Institute of High Performance Computing, 1 Science Park Road, #01-01 The Capricorn, Singapore Science Park II, Singapore 117528, and Department of Electrical Engineering, National University of Singapore, Lower Kent Ridge Road, Singapore 119260

Received December 14, 2001. Revised Manuscript Received April 22, 2002

Band gap energy and lattice constant of ternary semiconductors (ABC_2 chalcopyrites) are correlated to their chemical stoichiometrics and fundamental element properties of the constituents, through artificial intelligence techniques and sublattice model without prior knowledge of the electronic structures of individual compounds. New chalcopyrite semiconductors of $I-III-VI_2$ and $II-IV-V_2$ types are then predicted by using the developed model. Many potential semiconductors are selected from 270 possible new chalcopyrites by screening the element periodic table. The predicted band gap energy and lattice constant of the potential semiconductors fall in the range of 0.10−4.96 eV and 2.22−10.14 Å, respectively, which may offer a useful guideline in the exploration of new semiconductors for a wide range of applications by further experimental or first principles studies.

1. Introduction

Ternary ABC_2 chalcopyrites, including $I-III-VI_2$ and $II-IV-V_2$ compounds, form a large group of semiconductor materials with diverse optical, electrical, and structural properties. Great effort has been devoted to investigate these semiconductors for desirable properties.[1,2] It would be a formidable task, if not impossible, to explore all these materials solely by growing and experimentally measuring their physical properties. In our previous work,[3] we exploited the predictive power of a correlation technique to obtain the band gap energies (BEs) and lattice constants (LCs) for III−V and II−VI semiconductors, thus providing a guideline for future experimental analysis of interesting compounds. The correlation model was based on the chemical stoichiometrics and fundamental element properties of the constituents as well as the sublattice model, without relying on other additional knowledge of the compounds. In the present contribution, we extend this correlation technique to the prediction of band gap energies and lattice constants of ternary ABC_2 semiconductors.

The correlation technique used in the present work is an artificial intelligence expert system—APEX (Advanced Process Expert) which was developed in house on the IBM SP2 supercomputer of the Institute of High Performance Computing.[4] This tool includes the following components: data processing, collinearity check, feature reduction, pattern recognition, and neural network prediction. The first three components are integrated with pattern recognition to correlate the semiconductor properties to significant factors associated with the constituent element properties. Artificial neural network methods have the intrinsic ability to handle systems involving many unknown variables and relationships that are very complex and poorly understood. This is particularly useful in the present situation because the correlation of the semiconductor properties and the constituent element properties is very complicated in nature. A developed neural network model can be used to quantitatively predict the band gap energy and lattice constant.

A maximum number of 75 $I-III-VI_2$ compounds is formed from elements of groups IB, IIIB, and VIB. Similarly, 225 $II-IV-V_2$ compounds can be formed from elements of groups IIA, IIB, IVB, and VB. Among the total 300 compounds, band gap energies of 33 compounds and lattice constants of 37 compounds are reported in the literature. Other compounds are potential semiconductors to be studied, and their band gap energies and lattice constants are predicted in the present study. In the literature, the energy gaps and lattice constants of mixed chalcopyrites formed respectively by either two $I-III-VI_2$ or two $II-IV-V_2$ parent chalcopyrites are reported. To the authors' best knowledge, during the last few decades, studies of band gap energy and lattice constants of chalcopyrites are mainly focused on the known 37 compounds and their mixtures. Very little knowledge is available about the rest. Therefore, the present emphasis is placed on the study of the two important bulk properties of these new potential semiconductor materials.

This paper is organized as follows. We begin in section 2 with the definition of the correlation variables on the basis of chemical stoichiometrics and fundamental ele-

* To whom correspondence should be addressed. Email: wuping@ihpc.nus.edu.sg. Telephone: (65) 64191212. Fax: (65) 67780522.
† Institute of High Performance Computing.
‡ National University of Singapore.

(1) Jaffe, J. E.; Zunger, A. *Phys Rev B*: *Condens. Matter Mater. Phys.* **1984**, *29*, 1882.
(2) Petukhov, A. G.; Lambrecht, W. R. L.; Segall, B. *Phys Rev B*: *Condens. Matter Mater. Phys.* **1994**, *49*, 4549.
(3) Heng, K. L.; Chua, S. J.; Wu, P. *Chem. Mater.* **2000**, *12*, 1648.
(4) Heng, K. L.; Jin, H. M.; Li, Y.; Wu, P. *J. Mater. Chem.* **1999**, *9*, 837.

10.1021/cm0103996 CCC: $22.00
Published on Web 06/22/2002

ment properties as well as sublattice model. Next in section 3, we present the result of pattern recognition and neural network prediction of band gap energy and lattice constant for new ABC_2 compounds. In section 4, we discuss the confidence level of the model predictions and demonstrate that the band gap energy estimation for ternary systems is in agreement with a general tendency observed in binary systems. The estimated lattice constants are compared with the result from a simple empirical method. Finally in section 5, we make conclusions of the present work.

2. Computational Details

To correlate the semiconductor properties of ternary compounds to their constituent element properties and perform prediction of band gap energies and lattice constants, we follow the similar methodologies as used in the prediction of binary compounds.[3] First, the choice of element properties is made, followed by defining the model variables used for the prediction of the bulk properties. Next, pattern recognition and neural network prediction of band gap energies for all the new chalcopyrites are performed so as to narrow down the number of semiconductor candidates. Then, in the same manner as above, lattice constants of the selected compounds are predicted. The number of candidate semiconductors is, then, further reduced.

As mentioned previously, the entire set of chalcopyrites consists of two types of compounds, i.e., I−III−VI_2 and II−IV−V_2. For each type, the compound is assumed to form by the combination of three elements, each belonging to one of the following three groups: (type 1) I−III−VI_2 chalcopyrites; (group IB) Cu, Ag, Au; (group IIIB) B, Al, Ga, In, Tl; (group VIB) O, S, Se, Te, Po; (type 2) II−IV−V_2 chalcopyrites; (group IIA, IIB) Zn, Cd, Hg, Be, Mg, Ca, Sr, Ba, Ra; (group IVB) C, Si, Ge, Sn, Pb; (group VB) N, P, As, Sb, Bi.

Full combinations of type 1 form 75 (=3 × 5 × 5) compounds from elements of groups IB, IIIB, and VIB. Among which, 18 compounds have been reported in the literature with the band gap energies and 23 with lattice constants, and the remaining compounds are the candidates to be studied. Similarly, 225 (=9 × 5 × 5) compounds of type 2 are formed from elements of groups IIA, IIB, IVB, and VB. Fifteen of them have been reported in the literature with BE and 14 with LC values. The rest of the type 2 compounds are to be studied. Tables 1 and 2 list the compounds of types 1 and 2, respectively, whose BE and LC are available in the literature.[5]

The choice of element properties is similar to our previous study for binary semiconductors. Chemical element properties are generally classified into six groups,[6] namely, electrochemical factor, atomic number factor, cohesive energy factor, size factor, valence-electron factor, and Mendenleev number factor. Specially, the following chemical element properties are used to correlate band gap energy and lattice constant of the ternary semiconductors: Martynov−Batsanov electronegativity ($eV^{1/2}$), atomic number, melting temperature (K), Zunger pseudopotential radii sum (atomic units, au), valence electron, and Mendeleev number (MN) 1 to 4. For the first five properties, each of them is selected from the first five groups, respectively. The other four are from the last group. Although four sets of Mendeleev numbers, namely, MN1, MN2, MN3, and MN4, are available, it is found that the resulting property map coordinates (define later) of MN3 have the same values as that of the corresponding MN1, and similarly for MN2 and MN4. Therefore, MN3 and MN4 are excluded from our choice of the chemical element properties. Tables 3 and 4 summarize the constituent element properties for chalcopyrites of type 1 and type 2, respectively.[6,7]

The crystal structures of chalcopyrites ABC_2 can be found from the handbook.[8] A three-sublattice model is used for all the chalcopyrites. Lattice A is occupied by either group I (for type 1) or group II (for type 2) elements, lattice B by either group III (for type 1) or group IV (for type 2), and lattice C by either group VI (for type 1) or group V (for type 2). For each ABC_2

Table 1. Type 1 Chalcopyrites with Known Data from Literature[5]

no.	compound	BE (eV)	LC (Å)	Cu	Ag	Al	Ga	In	Tl	Fe	S	Se	Te
1	$CuAlS_2$	3.49	5.32	1	0	1	0	0	0	0	2	0	0
2	$CuAlSe_2$	2.65	5.606	1	0	1	0	0	0	0	0	2	0
3	$CuAlTe_2$	2.06	5.964	1	0	1	0	0	0	0	0	0	2
4	$CuGaS_2$	2.43	5.349	1	0	0	1	0	0	0	2	0	0
5	$CuGaSe_2$	1.68	5.607	1	0	0	1	0	0	0	0	2	0
6	$CuGaTe_2$	1.24	5.994	1	0	0	1	0	0	0	0	0	2
7	$CuInS_2$	1.53	5.517	1	0	0	0	1	0	0	2	0	0
8	$CuInSe_2$	1.04	5.773	1	0	0	0	1	0	0	0	2	0
9	$CuInTe_2$	1.06	6.167	1	0	0	0	1	0	0	0	0	2
10	$AgAlS_2$	3.13	5.695	0	1	1	0	0	0	0	2	0	0
11	$AgAlSe_2$	2.55	5.956	0	1	1	0	0	0	0	0	2	0
12	$AgAlTe_2$	2.27	6.296	0	1	1	0	0	0	0	0	0	2
13	$AgGaS_2$	2.64	5.743	0	1	0	1	0	0	0	2	0	0
14	$AgGaSe_2$	1.80	5.973	0	1	0	1	0	0	0	0	2	0
15	$AgGaTe_2$	1.32	6.283	0	1	0	1	0	0	0	0	0	2
16	$AgInS_2$	1.87	5.816	0	1	0	0	1	0	0	2	0	0
17	$AgInSe_2$	1.24	6.090	0	1	0	0	1	0	0	0	2	0
18	$AgInTe_2$	0.95	6.406	0	1	0	0	1	0	0	0	0	2
19	$CuTlS_2$	-	5.580	1	0	0	0	0	1	0	2	0	0
20	$CuTlSe_2$	-	5.832	1	0	0	0	0	1	0	0	2	0
21	$CuFeS_2$	-	5.25	1	0	0	0	0	0	1	2	0	0
22	$CuFeSe_2$	-	5.67	1	0	0	0	0	0	1	0	2	0
23	$AgTlSe_2$	-	9.70	0	1	0	0	0	1	0	0	2	0

Table 2. Type 2 Chalcopyrites with Known Data from Literature[5]

no.	compound	BE (eV)	LC (Å)	Mg	Zn	Cd	Si	Ge	Sn	N	P	As	Sb
1	$ZnSiP_2$	2.07	5.399	0	1	0	1	0	0	0	2	0	0
2	$ZnGeP_2$	1.99	5.465	0	1	0	0	1	0	0	2	0	0
3	$ZnSnP_2$	1.66	5.652	0	1	0	0	0	1	0	2	0	0
4	$ZnSiAs_2$	1.94	5.601	0	1	0	1	0	0	0	0	2	0
5	$ZnGeAs_2$	1.15	5.672	0	1	0	0	1	0	0	0	2	0
6	$ZnSnAs_2$	0.73	5.851	0	1	0	0	0	1	0	0	2	0
7	$CdSiP_2$	2.10	5.680	0	0	1	1	0	0	0	2	0	0
8	$CdGeP_2$	1.72	5.740	0	0	1	0	1	0	0	2	0	0
9	$CdSnP_2$	1.17	5.901	0	0	1	0	0	1	0	2	0	0
10	$CdSiAs_2$	1.55	5.885	0	0	1	1	0	0	0	0	2	0
11	$CdGeAs_2$	0.57	5.943	0	0	1	0	1	0	0	0	2	0
12	$CdSnAs_2$	0.26	6.093	0	0	1	0	0	1	0	0	2	0
13	$ZnSnSb_2$	0.70	6.275	0	1	0	0	0	1	0	0	0	2
14	$MgSiP_2$	2.82	5.718	1	0	0	1	0	0	0	2	0	0
15	$ZnGeN_2$	2.67	-	0	1	0	0	1	0	2	0	0	0

(5) Madelung, O.; Schulz, M.; Weiss, H. *Numerical Data and Functional Relationships in Science and Technology*; Springer-Verlag: Berlin, 1984; New Series. Group III: Crystal and Solid State Physics. Volume 17 Semiconductors. Subvolume h, Physics of Ternary Compounds.

(6) Villars, P.; Brandenburg, K.; Berndt, M.; et al. *J. Alloy Compd.* **2001**, *317−318*, 26.

(7) Villars, P.; Hulliger, F. *J. Less-Common Met.* **1987**, *132*, 289.

(8) Daams, J. L. C.; Villars, P.; van Vucht, J. H. N. *Atlas of Crystal Structure Types for Intermetallic Phases*; ASM International: Materials Park, OH, 1991; Vol. 3, p 4157.

Table 3. Basic Element Properties of Type 1 Compounds

element	electronegativity ($eV^{1/2}$)	atomic no.	melting temp (K)	pseudopotential radii sum (au)	valence electron (V)	Mendeleev no. 1	Mendeleev no. 2
Cu	1.08	29	1356	2.04	11	66	68
Ag	1.07	47	1235	2.375	11	67	67
Au	1.19	79	1337.6	2.66	11	68	66
B	1.90	5	2365	0.795	3	72	76
Al	1.64	13	933	1.675	3	73	75
Ga	1.70	31	303	1.695	3	74	74
In	1.63	49	430	2.05	3	75	73
Tl	1.69	81	577	2.235	3	76	72
Fe	1.67	26	1808	2.11	8	57	59
O	3.32	8	54.8	0.465	6	87	91
S	2.65	16	386	1.10	6	88	90
Se	2.54	34	490	1.285	6	89	89
Te	2.38	52	723	1.67	6	90	88
Po	2.40	84	527	1.90	6	91	87

Table 4. Basic Element Properties of Type 2 Compounds

element	electronegativity ($eV^{1/2}$)	atomic no.	melting temp (K)	pseudopotential radii sum (au)	valence electron (V)	Mendeleev no. 1	Mendeleev no. 2
Be	1.45	4	1560	1.08	2	7	12
Mg	1.31	12	922	2.03	2	8	11
Ca	1.17	20	1112	3.00	2	9	10
Sr	1.13	38	1042	3.21	2	10	9
Ba	1.08	56	1002	3.402	2	11	8
Ra	0.90	88	973	3.53	2	12	7
Zn	1.44	30	692	1.88	12	69	71
Cd	1.40	48	594	2.215	12	70	70
Hg	1.49	80	234	2.41	12	71	69
C	2.37	6	3825	0.64	4	77	81
Si	1.98	14	1683	1.42	4	78	80
Ge	1.99	32	1211	1.56	4	79	79
Sn	1.88	50	505	1.88	4	80	78
Pb	1.92	82	601	2.09	4	81	77
N	2.85	7	63	0.54	5	82	86
P	2.32	15	317	1.24	5	83	85
As	2.27	33	1090	1.415	5	84	84
Sb	2.14	51	904	1.765	5	85	83
Bi	2.14	83	544	1.997	5	86	82

compound, the lattice fractions of lattice A are summed up to unity and similarly for lattice B. For lattice C, the lattice fraction is summed up to unity and then multiplied by 2, which is equal to the stoichiometrics of C in the ABC_2 compound. The lattice fraction is defined as

$$Y_{A_i} = \frac{n_{A_i}}{\sum_{i=1}^{I} n_{A_i}} \quad (1)$$

$i = 1, ..., I$ ($I = 3$ for type 1 and $I = 9$ for type 2)

$$Y_{B_j} = \frac{n_{B_j}}{\sum_{j=1}^{J} n_{B_j}} \quad (2)$$

$j = 1, ..., J$ ($J = 5$ for both type 1 and type 2)

$$Y_{C_k} = 2\frac{n_{C_k}}{\sum_{k=1}^{K} n_{C_k}} \quad (3)$$

$k = 1, ..., K$ ($K = 5$ for both type 1 and type 2)

where n_{A_i}, n_{B_j}, and n_{C_k} are the number of moles of A_i, B_j, and C_k in lattice A, B, and C, respectively.

To build a property map, we use the similar definition as given in the binary system[3] where the coordinates are the functions of element properties and element compositions in the compounds. For an A_i−B_j atomic pair, the absolute difference between atom A_i and B_j in terms of the property P (e.g., electronegativity) is obtained as

$$Q_{A_iB_j} = |P_{A_i} - P_{B_j}| \quad (4)$$

$i = 1, ..., I; j = 1, ..., J$

Similarly, for A_i−C_k and B_j−C_k atom pairs, we define

$$Q_{A_iC_k} = |P_{A_i} - P_{C_k}| \quad (5)$$

$i = 1, ..., I; k = 1, ..., K$

$$Q_{B_jC_k} = |P_{B_j} - P_{C_k}| \quad (6)$$

$j = 1, ..., J; k = 1, ..., K$

Subsequently, in a three sublattice compound of ABC_2, a property map coordinate O_p with respect to element property P, is defined as

$$O_p = \sum_{A_i}\sum_{B_j} Y_{A_i} Y_{B_j} Q_{A_iB_j} + \sum_{A_i}\sum_{C_k} Y_{A_i} Y_{C_k} Q_{A_iC_k} + \sum_{B_j}\sum_{C_k} Y_{B_j} Y_{C_k} Q_{B_jC_k} \quad (7)$$

O_p defined in eq 7 is the overall element property difference of all nearest-neighbor atomic pairs in the compound. Coordinates O_1, O_2, O_3, O_4, O_5, O_6, and O_7 represent the differences of electronegativity, atomic

number, melting temperature, pseudopotential radii, valence electron, and Mendeleev numbers 1 and 2, respectively. The coordinates are used for pattern recognition and property prediction, which will be presented in the following sections.

3. Results

3.1. Prediction of Band Gap Energies for All New Chalcopyrites. In this section, we start with developing the optimal visual pattern of 2D projections of the calculated coordinates for band gap energies, followed by the verification with the data of mixed chalcopyrites. Then the model is applied to all the new chalcopyrite candidates to classify them. Subsequently, a neural network (NN) model is built using the same set of known data that is used for building the pattern. The BE values of all the possible candidates are then estimated by the NN model. Finally, these BE estimations are checked with the class type determined from the pattern recognition model. A compound is qualified for further prediction of lattice constants only if the BE predictions by both models are in agreement with each other.

3.1.1. Pattern Recognition of Band Gap Energies. On the basis of the values of the band gap energies, 33 known compounds listed in Tables 1 and 2 (18 of type 1 and 15 of type 2) are divided into two classes with the ratio of sample numbers as 1:1 in each class. The class is assigned to be 1 for those with BE < 1.7 eV and 2 for those BE ≥ 1.7 eV.

A correlation between band gap energy of ternary chalcopyrites and the constituent element properties is established by the APEX tool. In particular, the following techniques are employed: principal component analysis (PCA), partial least squares (PLS), modified Fisher discriminant analysis (MF), and linear mapping (LM). These techniques are used to search among the entire multiple-dimensional space of coordinates O_1, O_2, O_3, O_4, and O_5 and their nonlinear terms of O_i/O_j (25 variables in total), the most significant variables that are central to band gap energy in chalcopyrites. It is noted that in the study of energy gap pattern, the factor of the Mendeleev number, namely, MN1 and MN2, does not result in better classification patterns than without them. On the other hand, the Mendeleev number plays a role in the study of lattice constant, as will be shown in the later sections.

By visual examination, it is found that MF generates the best result where clear separation of the two classes is observed (Figure 1). In this figure, compounds of class 1 (lower BE values) are distributed on the right upper zone and class 2 (higher BE values) on the left lower zone. Only two data points ($CuInS_2$ and $CdSiAs_2$) do not fall in the class zones that they belong to, which indicates that they are exceptions to the model.

The MF algorithm reduces the number of variables from 25 to 10 that are central to the class pattern and governing band gap energy. It is noted that the shortlist of 10 variables listed in Table 5 consists of O_1, O_2, O_3, and O_4 and some of their nonlinear terms. In the MF

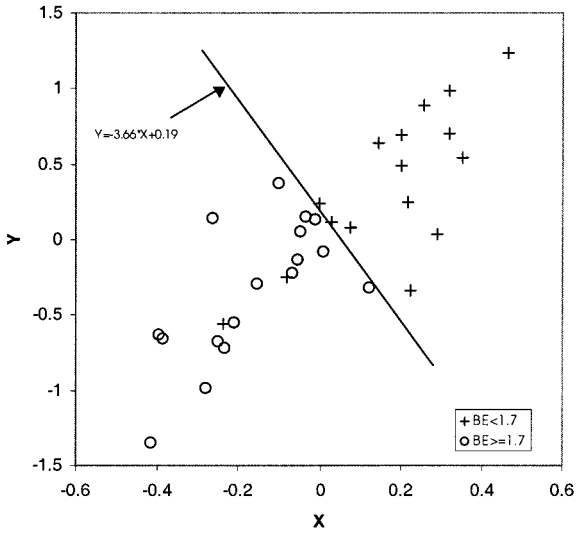

Figure 1. BE classification of known I−III−VI_2 and II−IV−V_2 chalcopyrites.

Table 5. Variables and Coefficients of MF for BE Classification

i	v_i	a_i	b_i
0		-5.92×10^{-1}	1.82
1	O2/O4	-1.93×10^{-2}	9.80×10^{-3}
2	O4/O1	9.28×10^{-1}	8.77×10^{-1}
3	O1	1.20×10^{-1}	2.18×10^{-1}
4	O2	-3.94×10^{-3}	-1.33×10^{-2}
5	O3	-2.90×10^{-4}	-5.29×10^{-4}
6	O4	-9.75×10^{-2}	-4.01×10^{-1}
7	O2/O3	-7.45×10^{-1}	−4.50
8	O1/O3	$-1.49 \times 10^{+2}$	$2.27 \times 10^{+2}$
9	O1/O4	6.24×10^{-1}	-3.31×10^{-1}
10	O2/O1	3.41×10^{-2}	1.56×10^{-2}

plot (Figure 1), the X and Y axes are given by the following equations

$$X = a_0 + \sum_{i=1}^{10} a_i v_i \tag{8}$$

$$Y = b_0 + \sum_{i=1}^{10} b_i v_i \tag{9}$$

where v_i, a_i, and b_i are the optimized variables and coefficients, respectively. They are listed in Table 5. It is noted that although the coefficients appear to vary significantly in magnitude, each term (the variable multiplying the coefficient) results in only a margin difference of 1 order of magnitude. This suggests that not a single factor solely determines the band gap energy of the compound. In comparison with the binary system where only two parameters are needed, the ternary system demonstrates a more complicated relationship between the band gap energy and the coordinates (variables). A separate line can be drawn for the BE classification (Figure 1) which is given by

$$Y = -3.66X + 0.19 \tag{10}$$

We further examine this pattern with the literature BE data of mixed chalcopyrites at different concentrations.[9] These mixed chalcopyrites include the following

(9) Shaukat, A. *J. Phys. Chem. Solids* **1990**, *51*, 1413.

Table 6. Band Gap Energy Values (eV) of Mixed Chalcopyrites at Different Concentrations

compound	$x = 0.2$	$x = 0.4$	$x = 0.6$	$x = 0.8$	compound	$x = 0.2$	$x = 0.4$	$x = 0.6$	$x = 0.8$
				Type 3					
$Ag_xCu_{1-x}AlS_2$	3.41	3.33	3.26	3.19	$Ag_xCu_{1-x}InSe_2$	1.07	1.11	1.15	1.19
$Ag_xCu_{1-x}GaS_2$	2.48	2.54	2.60	2.66	$Ag_xCu_{1-x}AlTe_2$	2.09	2.13	2.17	2.22
$Ag_xCu_{1-x}InS_2$	1.59	1.65	1.72	1.79	$Ag_xCu_{1-x}GaTe_2$	1.24	1.26	1.28	1.30
$Ag_xCu_{1-x}AlSe_2$	2.64	2.61	2.59	2.57	$Ag_xCu_{1-x}InTe_2$	0.95	0.95	0.95	0.95
$Ag_xCu_{1-x}GaSe_2$	1.70	1.73	1.76	1.79					
				Type 4					
$Zn_xCd_{1-x}SiP_2$	2.08	2.07	2.06	2.06	$Zn_xCd_{1-x}SiAs_2$	1.57	1.60	1.64	1.69
$Zn_xCd_{1-x}GeP_2$	1.76	1.81	1.86	1.92	$Zn_xCd_{1-x}GeAs_2$	0.67	0.78	0.90	1.02
$Zn_xCd_{1-x}SnP_2$	1.25	1.34	1.44	1.55	$Zn_xCd_{1-x}SnAs_2$	0.34	0.43	0.52	0.62
				Type 5					
$AgAl_xGa_{1-x}S_2$	2.80	2.88	2.96	3.04	$CuAl_xIn_{1-x}S_2$	1.91	2.29	2.69	3.09
$AgAl_xGa_{1-x}Se_2$	1.96	2.10	2.25	2.40	$CuAl_xIn_{1-x}Se_2$	1.35	1.67	2.00	2.33
$AgAl_xGa_{1-x}Te_2$	1.51	1.69	1.88	2.07	$CuAl_xIn_{1-x}Te_2$	1.17	1.38	1.60	1.83
$CuAl_xGa_{1-x}S_2$	2.63	2.84	3.05	3.27	$AgGa_xIn_{1-x}S_2$	2.01	2.17	2.33	2.51
$CuAl_xGa_{1-x}Se_2$	1.87	2.06	2.26	2.46	$AgGa_xIn_{1-x}Se_2$	1.34	1.44	1.56	1.69
$CuAl_xGa_{1-x}Te_2$	1.39	1.55	1.71	1.88	$AgGa_xIn_{1-x}Te_2$	1.01	1.07	1.15	1.23
$AgAl_xIn_{1-x}S_2$	2.11	2.35	2.61	2.87	$CuGa_xIn_{1-x}S_2$	1.63	1.81	2.01	2.21
$AgAl_xIn_{1-x}Se_2$	1.49	1.74	2.01	2.28	$CuGa_xIn_{1-x}Se_2$	1.15	1.26	1.39	1.53
$AgAl_xIn_{1-x}Te_2$	1.21	1.46	1.73	2.00	$CuGa_xIn_{1-x}Te_2$	0.99	1.03	1.09	1.15
				Type 6					
$ZnSi_xGe_{1-x}P_2$	1.99	1.99	2.01	2.03	$CdSi_xSn_{1-x}P_2$	1.35	1.53	1.71	1.90
$ZnSi_xGe_{1-x}As_2$	1.25	1.35	1.47	1.60	$CdSi_xSn_{1-x}As_2$	0.51	0.76	1.02	1.28
$CdSi_xGe_{1-x}P_2$	1.78	1.84	1.92	2.00	$ZnGe_xSn_{1-x}P_2$	1.70	1.75	1.81	1.89
$CdSi_xGe_{1-x}As_2$	0.75	0.93	1.13	1.33	$ZnGe_xSn_{1-x}As_2$	0.78	0.85	0.94	1.04
$ZnSi_xSn_{1-x}P_2$	1.73	1.81	1.89	1.98	$CdGe_xSn_{1-x}P_2$	1.25	1.34	1.45	1.58
$ZnSi_xSn_{1-x}As_2$	0.92	1.12	1.32	1.53	$CdGe_xSn_{1-x}As_2$	0.29	0.34	0.40	0.48
				Type 7					
$AgAl(S_xSe_{1-x})_2$	2.62	2.72	2.84	2.97	$CuAl(S_xTe_{1-x})_2$	2.27	2.52	2.80	3.13
$AgGa(S_xSe_{1-x})_2$	1.97	2.13	2.31	2.51	$CuGa(S_xTe_{1-x})_2$	1.39	1.59	1.83	2.11
$AgIn(S_xSe_{1-x})_2$	1.32	1.43	1.56	1.70	$CuIn(S_xTe_{1-x})_2$	1.00	1.07	1.19	1.34
$CuAl(S_xSe_{1-x})_2$	2.79	2.94	3.10	3.28	$AgAl(Se_xTe_{1-x})_2$	2.29	2.33	2.39	2.46
$CuGa(S_xSe_{1-x})_2$	1.79	1.92	2.07	2.24	$AgGa(Se_xTe_{1-x})_2$	1.39	1.47	1.58	1.69
$CuIn(S_xSe_{1-x})_2$	1.10	1.17	1.27	1.39	$AgIn(Se_xTe_{1-x})_2$	0.98	1.02	1.08	1.15
$AgAl(S_xTe_{1-x})_2$	2.37	2.50	2.67	2.88	$CuAl(Se_xTe_{1-x})_2$	2.15	2.25	2.37	2.51
$AgGa(S_xTe_{1-x})_2$	1.53	1.77	2.05	2.37	$CuGa(Se_xTe_{1-x})_2$	1.28	2.36	1.45	1.55
$AgIn(S_xTe_{1-x})_2$	1.07	1.21	1.39	1.61	$CuIn(Se_xTe_{1-x})_2$	0.94	0.94	0.96	0.99
				Type 8					
$ZnSi(P_xAs_{1-x})_2$	1.78	1.83	1.90	1.98	$CdSi(P_xAs_{1-x})_2$	1.63	1.73	1.84	1.96
$ZnGe(P_xAs_{1-x})_2$	1.29	1.45	1.62	1.80	$CdGe(P_xAs_{1-x})_2$	0.77	0.99	1.22	1.46
$ZnSn(P_xAs_{1-x})_2$	0.89	1.06	1.25	1.45	$CdSn(P_xAs_{1-x})_2$	0.42	0.59	0.77	0.96

types which contain 180 mixed $I-III-VI_2$ and 96 mixed $II-IV-V_2$ compounds, respectively (Table 6): (type 3) (Ia, Ib)-$III-VI_2$; (type 4) (IIa, IIb)-$IV-V_2$; (type 5) $I-$(IIIa, IIIb)-VI_2; (type 6) $II-$(IVa, IVb)-V_2; (type 7) $I-III-$(VIa,VIb)$_2$; (type 8) $II-IV-$(Va,Vb)$_2$.

It has been shown that the band gap energy of the mixed chalcopyrite varies almost linearly with the concentrations of the end ternaries. Therefore, it is expected that the pattern of mixed chalcopyrites should be consistent with that of their parents, i.e., the end ternaries. By simply replacing the corresponding element properties and the compositions, the abscissa and ordinate values are re-evaluated based on eqs 8 and 9 and the coefficients in Table 5, without performing pattern recognition. The resulting abscissa and ordinate values are plotted in Figure 2 where compounds associated with $CuInS_2$ and $CdSiAs_2$ have been removed. The two sets of compounds shown by crosses and the circles represent the two classes of the mixed chalcopyrites. The classes are determined as for the parent compounds, namely, class is 1 if BE < 1.7 eV and 2 if BE ≥ 1.7 eV. It can be seen that most mixed $I-III-VI_2$ chalcopyrites of class 1 are distributed on the upper right zone and similarly those of class 2 on the lower left zone. The distribution of the mixed compounds is

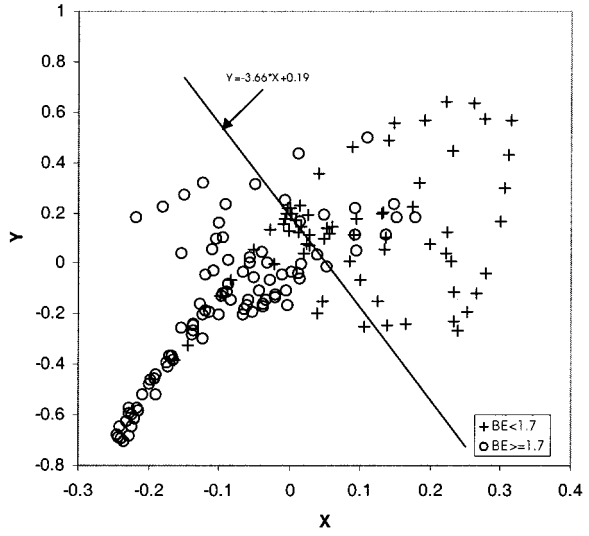

Figure 2. BE classification of mixed $I-III-VI_2$ compounds (types 3, 5, and 7).

similar as that of their parents in Figure 1. The data of the mixed $II-IV-V_2$ compounds are also treated in the same way. Similar behavior of this set of data is observed in Figure 3. These results verify the developed BE pattern for both parent and mixed chalcopyrites.

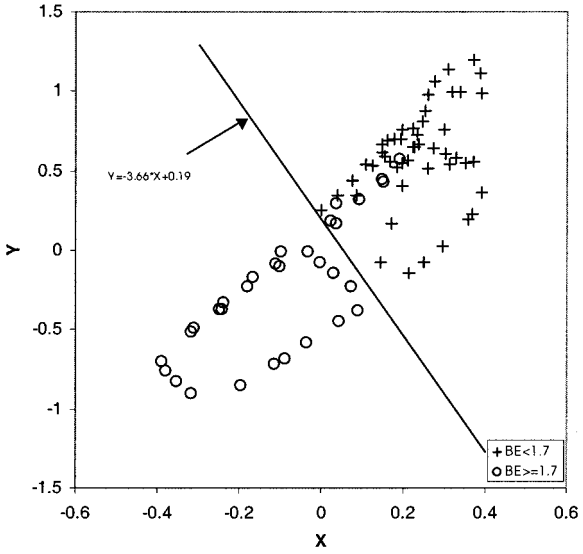

Figure 3. BE classification of mixed II−IV−V_2 compounds (types 4, 6, and 8).

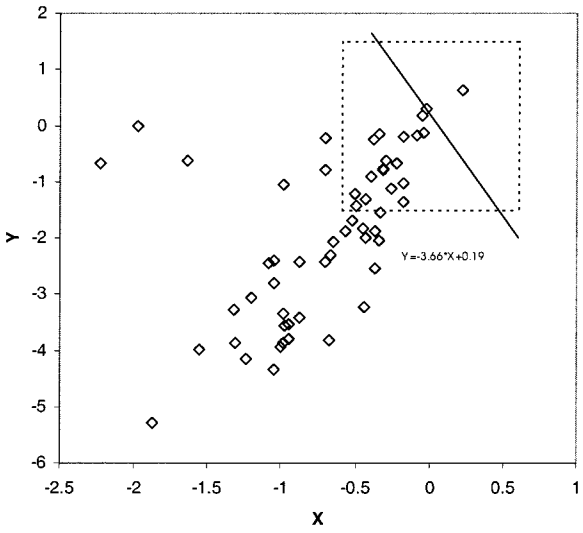

Figure 4. BE classification of new I−III−VI_2 compounds. The confident zone is located in the up rectangle.

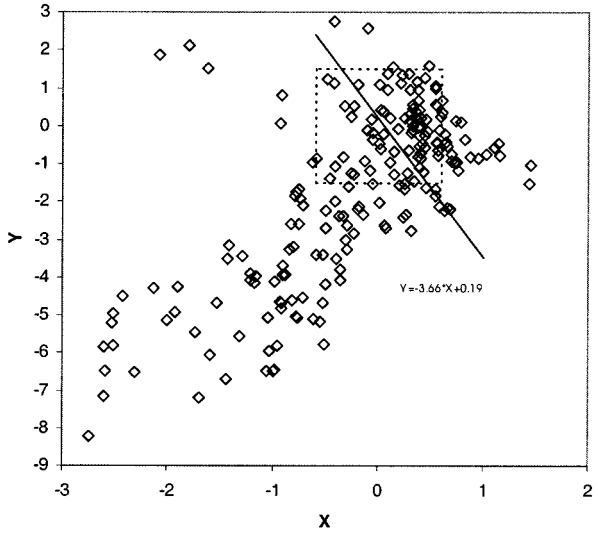

Figure 5. BE classification of new II−IV−V_2 compounds. The confident zone is located in the rectangle.

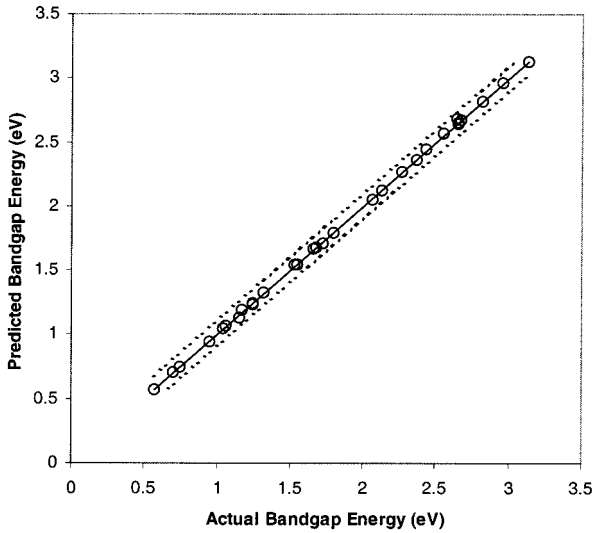

Figure 6. Comparison of actual and predicted band gap energy (training set).

Next, this pattern is used to classify all the new chalcopyrites. To do so, the *X* and *Y* axes for each new compound are needed to evaluate. With the same manner as the verification of mixed chalcopyrites, we replace the corresponding element properties and compositions of the new compound and then calculate the abscissa and ordinate values based on eqs 8 and 9 and the coefficients in Table 5. Each data point represents a new compound and is plotted together for classification.

The resultant distributions of all the new I−III−VI_2 and II−IV−V_2 chalcopyrites are depicted in Figures 4 and 5, respectively. The separate lines of eq 10 are placed on the two figures to divide the new compounds into two classes for each type. As indicated by the developed pattern, compounds distributing on the right-hand side of the line are determined as class 1 and on the left-hand side as class 2. Since the data of new compounds span a wider range than that of the known compounds, a zone of the known compounds of Figure 1 can be identified within the pattern plot of the new compounds. We call it the confident zone, which indicates the prediction is confident for those compounds within this zone. In contrast, outside the confident zone, we call it less confident zone where the extrapolation is applied for the prediction.

3.1.2. Neural Network Prediction of Band Gap Energies. The known BE data of chalcopyrites listed in Tables 1 and 2 are used to train the NN model by the Neural Network Model Builder (NNMB) of APEX. The predictors of the model are the significant variables selected from pattern recognition, i.e., the 10 variables listed in Table 5. The accuracy of the NN model is examined by comparing the actual with the predicted BE values of the known data. Randomly selected 90% of the samples are used for model training and the rest for model validation. The result of the comparison is shown in Figures 6 and 7 for the training and validation, respectively. As shown in the two figures, the uncertainties of the prediction of BE values fall in ±0.1 and ±0.5 (eV) for the training and validation data sets, respectively. It is noted that the choice of the samples for training and validation sets is randomly done by the algorithm routines. The samples of the validation set

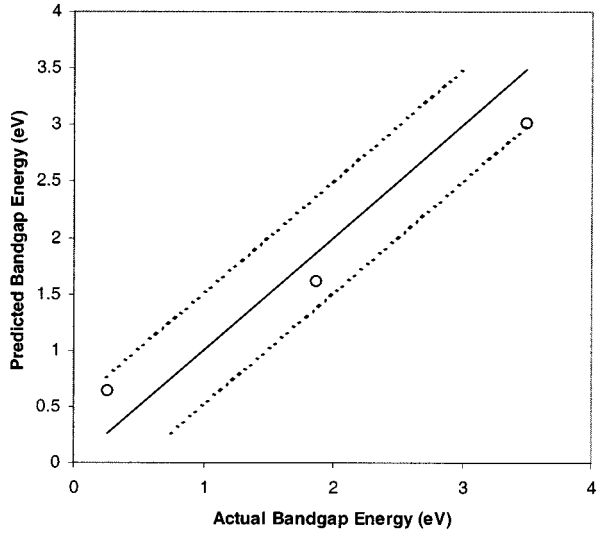

Figure 7. Comparison of actual and predicted band gap energy (validation set).

Table 7. Result of Estimated Band Gap Energies of New I−III−VI_2 Chalcopyrites

compound	estimated BE (eV)	class	compound	estimated BE (eV)	class
		In Confident Zone			
$CuAlO_2$	3.06	2	$AuAlPo_2$	2.58	2
$CuGaO_2$	3.20	2	$AuGaPo_2$	1.74	2
$CuTlSe_2$	1.95	2	$AuTlSe_2$	2.64	2
$AgTlSe_2$	1.78	2	$AuTlPo_2$	1.01	1
$AuAlTe_2$	2.47	2			
		In Less Confident Zone			
$CuBO_2$	3.88	2	$AgGaO_2$	2.40	2
$CuBSe_2$	1.96	2	$AgGaPo_2$	1.82	2
$CuBTe_2$	2.85	2	$AgInO_2$	2.18	2
$CuBPo_2$	2.55	2	$AgTlS_2$	2.52	2
$CuAlPo_2$	1.96	2	$AuBO_2$	2.47	2
$CuInO_2$	2.99	2	$AuBS_2$	4.08	2
$CuTlS_2$	2.83	2	$AuBSe_2$	4.10	2
$AgBO_2$	4.04	2	$AuBTe_2$	3.16	2
$AgBS_2$	3.49	2	$AuBPo_2$	3.62	2
$AgBSe_2$	3.26	2	$AuAlS_2$	3.07	2
$AgBTe_2$	2.86	2	$AuAlSe_2$	2.99	2
$AgBPo_2$	2.70	2	$AuGaS_2$	3.01	2
$AgAlO_2$	2.93	2	$AuInS_2$	2.79	2
$AgAlPo_2$	1.94	2	$AuTlS_2$	2.12	2

serve as only the validation purpose and do not involve model building. It is found in the present case that three compounds, namely, $CuAlS_2$, $CdSnAs_2$, and $AgInS_2$, are picked up for the validation. The BE values (eV) of the three compounds in the literature[5] are 3.49, 0.26, and 1.87, respectively, while their predicted values are 3.02, 0.64, and 1.62, respectively.

Next, the NN model is applied to all new I−III−VI_2 and II−IV−V_2 chalcopyrites for the estimation of the band gap energies. The estimated BE values are checked and compared with the determined class for each compound, i.e., whether the BE ≥ 1.7 eV or BE < 1.7 eV. Only those with BE values in agreement with their classes are regarded as acceptable and selected for subsequent study. The resultant selection of the I−III−VI_2 compounds is listed in Table 7 where the compounds are grouped based on their location in the confident or less confident zones. A similar selection for II−IV−V_2 compounds can also be made in the same way.

3.2. Prediction of Lattice Constants for the Selected Candidates. In the same manner as dealing

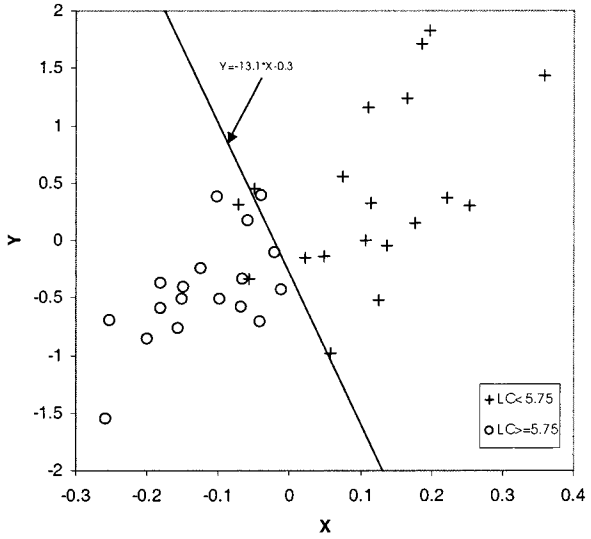

Figure 8. LC classification of known I−III−VI_2 and II−IV−V_2 chalcopyrites.

Table 8. Variables and Coefficients of MF for LC Classification

i	v_i	a_i	b_i
0		−1.75	−3.16
1	O3/O4	7.91×10^{-5}	6.85×10^{-4}
2	O6/O3	3.17×10	-3.69×10
3	O3/O2	3.06×10^{-3}	1.58×10^{-2}
4	O2	-5.26×10^{-3}	6.65×10^{-3}
5	O3	-2.55×10^{-5}	1.09×10^{-4}
6	O4	1.90×10^{-1}	1.80×10^{-1}
7	O3/O6	-6.92×10^{-3}	6.26×10^{-3}
8	O4/O3	$-5.80 \times 10^{+2}$	9.30×10
9	O2/O3	1.72×10	3.04
10	O1/O3	$-8.83 \times 10^{+2}$	$1.42 \times 10^{+2}$
11	O1/O6	3.88×10	7.83

with the band gap energies, pattern recognition in terms of lattice constant is carried out and then verified by mixed chalcopyrites. The selected candidates from BE prediction are then classified based on the LC pattern so as to determine their LC classes. Subsequently, a neural network model in terms of lattice constant is built and used for the prediction of LC for all selected chalcopyrites. The final selection is made on the chalcopyrites whose estimated LC values are in agreement with their classes determined previously by pattern recognition.

3.2.1. Pattern Recognition of Lattice Constants. In this section, the visual patterns of the 2D projections of lattice constant are developed using the LC data of the known 37 chalcopyrites listed in Tables 1 and 2. The class is defined as 1 if LC< 5.75 (Å) and 2 if LC ≥ 5.75 (Å). For convenience, the classification is so determined that both classes have the equal numbers of compounds for each type of chalcopyrite.

Pattern recognition is performed for the data set with two types of chalcopyrites together and using all the factors (O_1−O_7). The best pattern is found from MF projection with 11 significant variables that govern the 2D pattern. We note that the 11 significant variables (listed in Table 8) cover all the factors except O_5 and O_7, which are automatically excluded by MF algorithm when searching among all the coordinate space. The projection of the best separation is shown in Figure 8 and the abscissa and ordinate in the figure are deter-

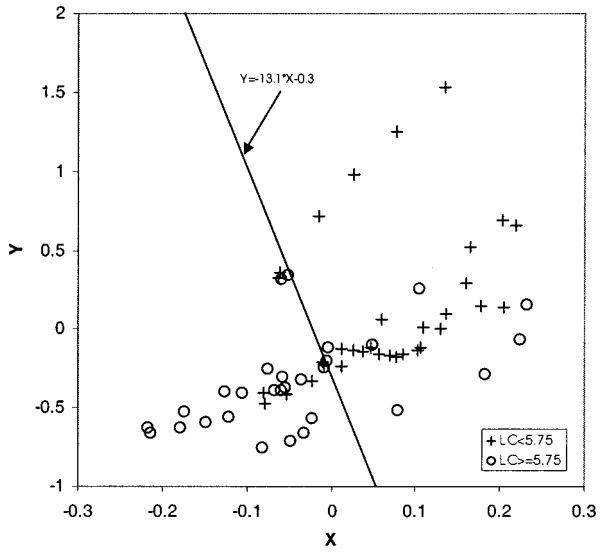

Figure 9. LC classification of mixed I−III−VI_2 and II−IV−V_2 chalcopyrites.

Table 9. Lattice Constant Values (Å) of Mixed Chalcopyrites at Different Concentrations

compounds	$x = 0.2$	$x = 0.4$	$x = 0.6$	$x = 0.8$
		I−III−VI_2		
$Cu_{1-x}Ag_xAlS_2$	5.4			
$Cu_{1-x}Ag_xAlSe_2$	5.68			
$Cu_{1-x}Ag_xGaS_2$	5.4	5.48		
$Cu_{1-x}Ag_xGaSe_2$	5.68	5.77		
$Cu_{1-x}Ag_xInS_2$	5.58	5.64	5.70	5.76
$Cu_{1-x}Ag_xInSe_2$	5.83	5.90	5.97	6.03
$CuAl(S_xSe_{1-x})_2$	5.57	5.50	5.45	5.38
$AgAl(S_xSe_{1-x})_2$	5.93	5.84	5.82	5.79
$CuGa(S_xSe_{1-x})_2$	5.55	5.50	5.46	5.39
$AgGa(S_xSe_{1-x})_2$	5.92	5.88	5.83	5.79
$CuIn(S_xSe_{1-x})_2$	5.70	5.66	5.60	5.55
$AgIn(S_xSe_{1-x})_2$	6.04	6.0	5.95	5.91
$CuGa_{1-x}In_xS_2$	5.39	5.43	5.46	5.49
$CuGa_{1-x}In_xSe_2$	5.65	5.67	5.72	5.75
		II−IV−V_2		
$Zn_{1-x}Cd_xSiAs_2$	5.66	5.72	5.77	5.84
$CdSn(P_xAs_{1-x})_2$	6.05	6.01	5.97	5.93
$CdGe_{1-x}Sn_xP_2$	5.77	5.81	5.84	5.87
$ZnSi(As_xP_{1-x})_2$	5.56	5.515	5.475	5.44

mined by the correlation analogous with eqs 8 and 9 where the variables and coefficients are listed in Table 8. Similar to the band gap energy pattern, not a single factor solely attributes significantly to the abscissa and ordinate values. The separation line for the pattern can be found as

$$Y = -13.1X - 0.3 \qquad (11)$$

It can be seen in Figure 8 that three data points are exceptions of the pattern. The three compounds are $CuAlSe_2$, $ZnGeAs_2$, and $CdSiAs_2$. The pattern is verified with the mixed chalcopyrites whose LC values vary with the composition[5] (Table 9). The abscissa and ordinate of the mixed chalcopyrites are computed by replacing their element properties and compositions and using the model variables and coefficients listed in Table 8, without performing pattern recognition. The resulting abscissa and ordinate are plotted in Figure 9 and the same separation line (eq 11) is also placed on this figure.

It is observed in Figure 9 that most of the mixed chalcopyrites distribute on the zones consistent with their class definition. But some of them violate the

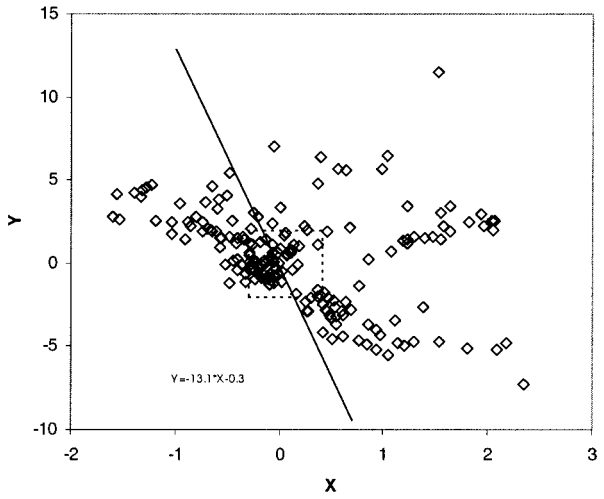

Figure 10. LC classification of new I−III−VI_2 and II−IV−V_2 candidates. The rectangle depicts the confident zone.

classes they belong to, i.e., those with low class (LC < 5.75) distribute in the zone of high class (LC ≥ 5.75) and those with high class fall in the low class zone.

A possible reason of the exception may be due to the accuracy of the model. It is found that the lattice constants of the compounds that distribute in the wrong zones are in the range of 5.60−5.87 (Å), while the uncertainty of the model is ±0.4 (Å) as will be shown in the later section. Therefore, the violation of the distribution of these compounds may be reversed if the model uncertainty is further reduced.

Next, the mapping of the LC classes of the candidates selected from BE estimation is performed. In the same manner as in determining the BE classes, abscissa and ordinate of the selected chalcopyrites are evaluated based on the correlation of the LC patterns whose variables and coefficients are listed in Table 8. The values of the variables are computed using the constituent element properties and compositions of the selected chalcopyrites. No pattern search is performed. The resulting distribution of LC is shown in Figure 10 for both the I−III−VI_2 and II−IV−V_2 candidates. The separate line of eq 11 divides the compounds into two classes. The compounds of lower LC class are determined for those locating on the right-hand side of the line and higher LC class on the left-hand side. The confident and less confident zones are determined by comparing the data ranges in their parent pattern. Table 10 shows the resulting classes of the I−III−VI_2 and indicates whether they fall in/out of the confident zone. The determination of the classes of II−IV−V_2 candidates is also made with the same fashion.

3.2.2. Neural Network Prediction of Lattice Constants. Neural network models for the prediction of lattice constants are trained by using the same data set of the 37 chalcopyrites listed in Tables 1 and 2 for LC pattern recognition. As mentioned previously, both I−III−VI_2 and II−IV−V_2 compounds are used for training a single model where randomly selected 90% of data are used for training and the rest for validation. The comparison of the predicted and actual LC values is shown in Figures 11 and 12 with the estimated uncertainty of ±0.4 Å and ±0.1 Å for the training and validation set, respectively.

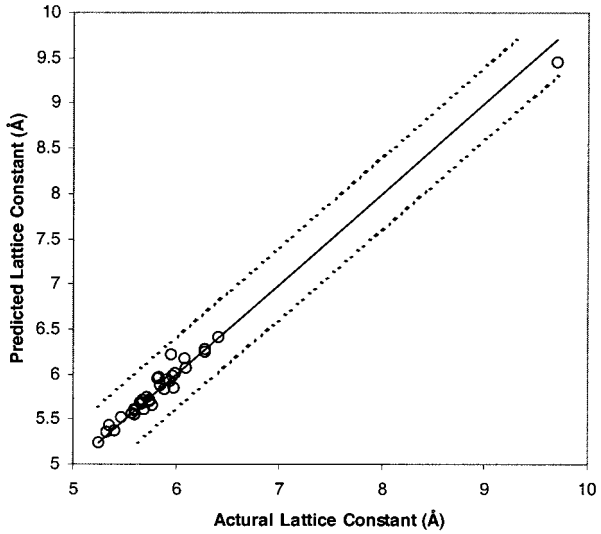

Figure 11. Comparison of actual and predicted lattice constant of I−III−VI_2 and II−IV−V_2 chalcopyrites (training set).

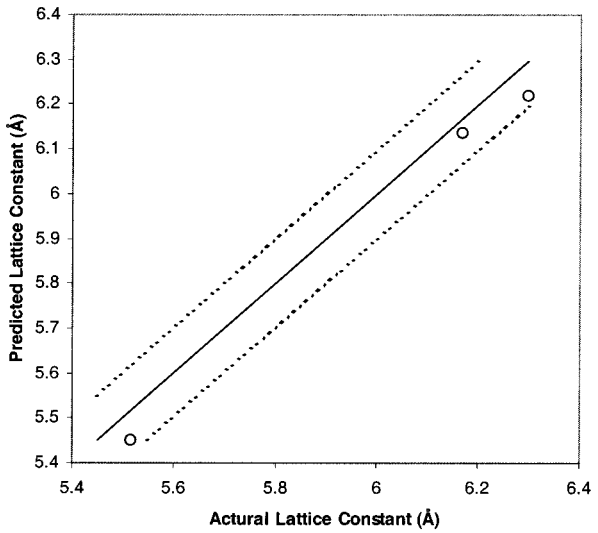

Figure 12. Comparison of actual and predicted lattice constant of I−III−VI_2 and II−IV−V_2 chalcopyrites (validation set).

Table 10. Result of Estimated LC for I−III−VI_2 Chalcopyrites

compound	LC class	LC zone	estimated LC (Å)	selection
		In Confident BE Zone		
$CuAlO_2$	1	out	3.1333	yes
$CuGaO_2$	1	out	2.3071	yes
$CuTlSe_2$	2	in	3.0023	
$AgTlSe_2$	2	in	3.6164	
$AuAlTe_2$	2	in	4.3862	
$AuAlPo_2$	2	in	3.5936	
$AuGaPo_2$	2	in	3.4157	
$AuTlSe_2$	2	in	4.3895	
$AuTlPo_2$	2	in	1.6895	
		In Less Confident BE Zone		
$CuBO_2$	1	out	4.7192	yes
$CuBSe_2$	1	in	4.9564	yes
$CuBTe_2$	2	out	5.1334	
$CuBPo_2$	2	out	4.8681	
$CuAlPo_2$	1	in	4.3881	yes
$CuInO_2$	1	out	3.3533	yes
$CuTlS_2$	1	in	3.0471	yes
$AgBO_2$	1	out	4.5691	yes
$AgBS_2$	1	in	4.7984	yes
$AgBSe_2$	1	in	4.9802	yes
$AgBTe_2$	2	in	5.2686	
$AgBPo_2$	2	out	5.0348	
$AgAlO_2$	1	out	3.9431	yes
$AgAlPo_2$	2	in	3.0299	
$AgGaO_2$	1	out	2.7096	yes
$AgGaPo_2$	2	in	2.7172	
$AgInO_2$	1	out	3.4877	yes
$AgTlS_2$	1	in	3.5845	yes
$AuBO_2$	1	out	4.375	yes
$AuBS_2$	1	out	5.266	yes
$AuBSe_2$	2	out	5.2923	
$AuBTe_2$	2	out	5.3262	
$AuBPo_2$	2	out	5.2845	
$AuAlS_2$	1	in	4.8521	yes
$AuAlSe_2$	2	in	4.234	
$AuGaS_2$	1	in	4.5258	yes
$AuInS_2$	1	in	4.6207	yes
$AuTlS_2$	1	in	5.0866	yes

Similarly as in building the NN model for energy gap, the choice of the samples for the training and validation sets of the LC model is done randomly by the algorithm, and the samples of the validation set are not used for the model building. In the present case, $CuInTe_2$, $AgAlTe_2$, and $CuInS_2$ are randomly picked up for the validation. The LC values (Å) of the three compounds in the literature[5] are 6.167, 6.296, and 5.517, in comparison with the predicted LC of 6.136, 6.220, and 5.449, respectively. In addition, the NN model is further examined with $AgTlTe_2$, a compound that is included neither in the training set nor in the validation set. The LC of $AgTlTe_2$ according to literature[5] is 3.92 Å and is estimated as 3.927 Å by the NN model.

Table 11. The Properties of the Final Selection of I−III−VI_2 Chalcopyrites

compound	estimate BE (eV)	estimate LC (Å)	BE zone	LC zone	LC[a] (Å)
$CuBO_2$	3.88	4.7192	out	out	4.1031
$CuBSe_2$	1.96	4.9564	out	in	5.3837
CuAlO2	3.06	3.1333	in	out	4.3977
$CuAlPo_2$	1.96	4.3881	out	in	
$CuGaO_2$	3.20	2.3071	in	out	
$CuInO_2$	2.99	3.3533	out	out	4.4795
$CuTlS_2$	2.83	3.0471	out	in	
$AgBO_2$	4.04	4.5691	out	out	4.3343
$AgBS_2$	3.49	4.7984	out	in	5.4009
$AgBSe_2$	3.26	4.9802	out	in	5.6322
$AgAlO_2$	2.93	3.9431	out	out	4.6939
$AgGaO_2$	2.40	2.7096	out	out	4.6901
$AgInO_2$	2.18	3.4877	out	out	
$AgTlS_2$	2.52	3.5845	out	in	
$AuBO_2$	2.47	4.375	out	out	4.3764
$AuBS_2$	4.08	5.266	out	out	5.4465
$AuAlS_2$	3.07	4.8521	out	in	5.7888
$AuGaS_2$	3.01	4.5258	out	in	5.785
$AuInS_2$	2.79	4.6207	out	in	5.9075
$AuTlS_2$	2.12	5.0866	out	in	

[a] LC is estimated by Abrahams−Bernstein equations.

The trained NN model is then used to estimate the LC values for the selected new compounds such as those listed in Table 7. The resulting LC values of I−III−VI_2 type are presented in Table 10 and then checked with their corresponding LC classes determined in the previous section. For those chalcopyrites of LC values in agreement with their class definition, namely, class is 1 if LC < 5.75 and class is 2 if LC ≥ 5.75, we mark them with "Yes" under the column of "Selection". The final 20 candidates of the new I−III−VI_2 chalcopyrites are summarized in Table 11. The selection of the II−IV−V_2 chalcopyrites can also be made in the same manner.

4. Discussions

4.1. Confidence of the Estimation. In the present work, the BE and LC values are estimated using artificial neural network method which usually generates reliable results within the range of the training set. For those candidates in Table 11 marked "in" under the column "BE zone" or "LC zone", we are more confident with the corresponding estimation. In contrast, we are less confident with the values if there are marked with "out". Obviously, the BE or LC values of some candidates do not fall in the confident zone as we extrapolate the calculation to the entire array of chemical elements in the periodic table. Nevertheless, the result still offers a general trend of the estimated semiconductor properties. As shown in Table 11, all the LC values of the selected I−III−VI_2 compounds are in the range of 2.307−5.266 Å, and BE values vary from 1.96 to 4.08 eV, which may be used for various applications.

4.2. Overview of the Estimation of Band Gap Energy. It is found that in the binary system there is a general tendency for an increasing energy gap with increasing ionic bonding.[10] For instance, I−VII compounds exhibit a larger ionicity than that of II−VI compounds, and further, II−VI compounds have larger ionicity than that of III−V compounds. Consequently, I−VII compounds have the largest BE values, followed by II−VI compounds. And III−V compounds have the smallest BE values among these three types of binary compounds. However, exceptions also exist, e.g., the energy gap of ZnS is larger than that of CuCl, but its ionicity is smaller than that of CuCl.

In a family of binary compounds, those comprising elements of lower atomic numbers are more ionic than those of higher atomic numbers, thus the energy gaps of the former are larger than that of the later. For instance, the energy gaps of GaP, GaAs, and GaSb are 2.350, 1.519, and 0.811 eV, respectively.[10] The present result of band gap energies of the ternary system (Table 11) also exhibits such tendency as the binary system.

The following lists compare the band gap energies of family members appearing in Table 11. It can be seen that most series are in agreement with the general trend. With the exception of series that are in italic, it tends to be difficult to explain due to more complicated structure of a ternary chalcopyrite.

$CuBO_2$ (3.88) > $CuBSe_2$ (1.96)

$CuAlO_2$ (3.06) > $CuAlPo_2$ (1.96)

$AgBO_2$ (4.04) > $AgBS_2$ (3.49) > $AgBSe_2$ (3.26)

$AuBO_2$ (2.47) < $AuBS_2$ (4.08)

$CuBO_2$ (3.88) > *$CuAlO_2$ (3.06) < $CuGaO_2$ (3.20)* > $CuInO_2$ (2.99)

$AgBO_2$ (4.04) > $AgAlO_2$ (2.93) > $AgGaO_2$ (2.40) > $AgInO_2$ (2.18)

$AgBS_2$ (3.49) > $AgTlS_2$ (2.52)

$AuBS_2$ (4.08) > $AuAlS_2$ (3.07) > $AuGaS_2$ (3.01) > $AuInS_2$ (2.79) > $AuTlS_2$ (2.12)

$CuBO_2$ (3.88) < $AgBO_2$ (4.04) > $AuBO_2$ (2.47)

$CuBSe_2$ (1.96) < $AgBSe_2$ (3.26)

$CuAlO_2$ (3.06) > $AgAlO_2$ (2.93)

$CuGaO_2$ (3.20) > $AgGaO_2$ (2.40)

$CuInO_2$ (2.99) > $AgInO_2$ (2.18)

$CuTlS_2$ (2.83) > $AgTlS_2$ (2.52) > $AuTlS_2$ (2.12)

Table 12. Tetrahedral Covalent Radii

Be			B	C	N	O
0.975			0.853	0.774	0.719	0.678
Mg			Al	Si	P	S
1.301			1.230	1.173	1.128	1.127
Ca	Cu	Zn	Ga	Ge	As	Se
1.333	1.225	1.225	1.225	1.225	1.225	1.225
Sr	Ag	Cd	In	Sn	Sb	Te
1.689	1.405	1.405	1.405	1.405	1.405	1.405
Ba[a]	Au[a]	Hg[a]		Pb[a]	Bi[a]	
1.98	1.44	1.49		1.47	1.46	

[a] Covalent radii from Suresh and Koga (2001).

4.3. Comparison of the Estimated Lattice Constants. To verify the estimation of the lattice constants of the new chalcopyrites, we employ the Abrahams−Bernstein equations[11] to approximate the lattice constants of the list of compounds. This simple method uses tetrahedral covalent radii of elements (Table 12) from Van Vechten and Phillips[12] and Suresh and Koga.[13]

The estimated lattice constants of the Abrahams−Bernstein analysis are listed in Table 11, in comparison with the prediction from our present work. It is found that the differences of the lattice constants estimated from the two methods vary between 0.0014 and 1.98 Å, although the estimation of Abrahams−Bernstein analysis is not very accurate.[14] Nevertheless, they can serve as a general check to our estimations.

5. Summary

In conclusion, the correlation technique used in the present work has served as an efficient tool to investigate semiconductor properties from a large number of compounds and narrow down the number of potential candidates. It would be an extreme time-consuming task if such exploration is done by experiments or first principle calculation. Although the present result may not be very accurate, it can offer a general guideline for further studies leading to the design of new semiconductors for specific applications.

CM0103996

(10) Grahn, H. T. *Introduction to Semiconductor Physics*; World Scientific Publishing Co. Pte. Ltd.: Singapore, 1999; p 7.

(11) Shay, J. L.; Wernick, J. H. *Ternary Chalcopyrite Semiconductors*: *Growth, Electronic Properties, and Applications*; Pergamon Press: Oxford, 1975.

(12) Van Vechten, J. A.; Phillips, J. C. *Phys Rev B*: *Condens. Matter Mater. Phys.* **1970**, *2*, 2160.

(13) Suresh, C. H.; Koga, N. *J. Phys. Chem. A* **2001**, *105*, 5940.

(14) Wang, T.; Moll, N. *Phys. Rev. B* **2000**, *63*, 035306.