PHYSICAL REVIEW A VOLUME 44, NUMBER 1 1 JULY 1991

Classification of Cm II and Pu I energy levels using counterpropagation neural networks

Keith L. Peterson
Division of Science and Mathematics, Wesleyan College, Macon, Georgia 31297
(Received 27 June 1990; revised manuscript received 15 March 1991)

Two different types of counterpropagation neural networks are applied to the problem of classifying unknown Cm II and Pu I energy levels according to their electronic configurations. Four features—energy level, angular momentum, g factor, and isotope shift—are used to describe each level. Both types of networks are trained at the 100% level for even-parity levels of Cm II, while for odd-parity levels one type is trained at 100% and the other at 97.5%. Performance on test sets is not as good, ranging from 80.0% to 93.3%. The two network types are trained at 98.6% and 90.5%, respectively, for even-parity levels of Pu I, and at 98.6% and 91.6% for odd-parity levels. Test-set performances range from 64.0% to 77.7%. Classifications for 159 previously unclassified Cm II levels and 485 previously unclassified Pu I levels are also obtained. Qualitative analysis of network-training results reveals several characteristics of those levels whose configurations the networks fail to learn, the most prevalent of which is an isotope-shift value that is consistent with two or more configurations. Rationalizations are also given for some of the other levels the networks misclassify. These qualitative observations should form a basis for understanding and precisely interpreting the performance of counterpropagation neural networks when applied to atomic systems.

I. INTRODUCTION

In a previous paper we have demonstrated the utility of counterpropagation neural networks as a tool in assigning configurations to the atomic energy levels of Cm I [1]. In this paper we apply the same types of networks to the classification of Cm II and Pu I energy levels. Our motivation for doing this is twofold: (1) we wish to characterize further the application of counterpropagation neural networks to atomic classification problems and (2) pattern-recognition methods (such as those used in classifying Cm I [2], U I [3], and U II [4] energy levels) failed miserably when applied to Cm II [5] and Pu I [6]. We will make several observations concerning the performance of the networks when applied to these systems, our goal being a better qualitative understanding of the networks than was established in Ref. [1].

For an introduction to neural networks and a detailed explanation of how counterpropagation neural nets are applied to the classification of atomic energy levels, the reader is referred to Ref. [1] and references therein. However, in order to make the present paper relatively self-contained we will, in the remainder of this section, briefly summarize the application of networks to the classification problem.

A neural network is provided with a set of inputs (a set of pattern vectors, each vector describing a given energy level with four features—energy, angular momentum J, g factor, and isotope shift, Δ_{IS}) and a set of desired outputs. Each output represents a given category (i.e., electronic configuration) to which the corresponding input belongs. Upon repeatedly presenting the input and output sets to the network, the weights between processing elements (PE's) are adjusted so that the network generates the correct output when given a certain input. After this training or learning phase the network is provided with an input whose corresponding output is unknown (i.e., an unclassified energy level). The network transforms the input into a predicted output (category or electronic configuration prediction). The category prediction process is referred to as recall. Recall can also be used with inputs whose corresponding outputs are known in order to determine the training level of the network.

Uniflow counterpropagation networks (see Ref. [1] and references therein) in the present application consist of four layers of PE's. The input layer contains four PE's—one PE for each feature—and acts as a buffer for the first hidden layer. This is a normalizing layer containing one more PE than the input layer. It ensures that every input vector has the same length. The second hidden layer is a competitive layer that acts as a nearest-neighbor classifier and contains as many PE's as there are inputs (energy levels). These PE's learn with the Kohonen learning rule. During training each PE competes with others in the layer and is equally likely to win for any randomly chosen input. For a given input, however, only one PE in this layer wins. The output of this winning PE serves as an input to the output layer which consists of one PE for each possible category (electronic configuration). The output layer uses the Widrow-Hoff learning rule to decode the output of the competitive layer into a predicted output. The normalizing layer is fully interconnected to the competitive layer which is in turn fully connected to the output layer.

A PE in the competitive layer of a uniflow network may take responsibility for two or more training inputs which belong to different categories. This may result in an ambiguous output for any inputs which activate this PE. A category-learning counterpropagation network solves this problem by preconditioning PE's in the competitive layer to learn only about a particular category. All of the above comments concerning uniflow networks also apply to category-learning networks except that each PE in the competitive layer is connected to only one PE in the output layer. Therefore there is no learning rule for the output layer.

II. APPLICATION TO Cm II ENERGY LEVELS

We now discuss the application of the two types of counterpropagation neural networks to the classification of Cm II energy levels. Data is taken from Worden, Conway, and Blaise [7]. Only energy levels for which four features (energy, J, g, and Δ_{IS}) are available are considered. This leaves us with two data sets, one consisting of 82 odd-parity levels representing four configurations ($5f^7 7s^2$, $5f^7 6d 7s$, $5f^7 6d^2$, and $5f^8 7p$) and one consisting of 74 even-parity levels representing five configurations ($5f^8 7s$, $5f^8 6d$, $5f^7 7s 7p$, $5f^7 6d 7p$, and $5f^7 6d 7p + 5f^8 6d$). The networks constructed for analyzing the odd-parity data set consist of an input layer of four PE's, a normalizing layer of five PE's, a competitive layer of 82 PE's, and an output layer of four PE's. Networks for the even-parity data set are similar except that they contain 74 PE's in the competitive layer and five PE's in the output layer. All networks were run on an IBM XT (with an 8088 chip, 640 k memory, 10-Mbyte hard disk, and CGA display) using the Neural Works Professional II™ software package [8]. All networks were trained on 9000 passes through an entire data set. As in our previous paper on Cm I, scaling of data was necessary in order for the networks to function properly. The data of Ref. [7] (with energy and Δ_{IS} expressed in cm^{-1}) were scaled by dividing energy by 100 000, and J and g by 10, leaving Δ_{IS} the same.

The training results of various networks are summarized in Table I. As in Ref. [1] we have trained networks with all classified levels and with "training sets." The even-parity and odd-parity training sets are formed by removing 15 classified levels from the original 74 even levels and 17 classified levels from the original 82 odd levels, respectively. The 15- and 17-member data sets are referred to at test sets. We then trained a network (with 59 PE's in the competitive layer) with the 59-member even-parity training set, and a network (with 65 PE's in the competitive layer) with the 65-member odd-parity training set. Each network was trained with 9000 passes through their respective training sets. Each member of the test set was then recalled through the trained network and the resulting outputs compared to the known outputs. The test-set results are generally not as good as the training results. However, they are quite respectable and are comparable to results obtained for Cm I [1]. In addition, they are far superior to pattern recognition results on Cm II [5].

We now make some observations which rationalize some of the mistakes reflected in Table I. Our motive for doing this is to obtain a better qualitative understanding and characterization of network performance; for example, we would like to begin to be able to identify those cases or levels where a network is likely to fail and those where it is likely to succeed. Ultimately (in later work), we would like to be able to relate input-output weights of the various PE's comprising the network to the performance of the network. Hopefully the weights of the input and normalizing layer PE's will give us insight into the relative importance of each of the four features in determining the classification of a given level.

A. Category-learning network; even-parity levels

The category-learning network misclassifies two levels in the even-parity training set. These two levels with their actual and predicted configurations are

$$35\,456.915,\quad 5.5,\quad 1.367,\quad -1.042$$

$$5f^7 6d 7p \text{ classified as } 5f^7 6d 7p + 5f^8 6d$$

and

$$35\,556.240,\quad 3.5,\quad 1.858,\quad -0.743$$

$$5f^7 6d 7p \text{ classified as } 5f^7 7s 7p\ .$$

[The ordered quartet of numbers corresponds to E (cm^{-1}), J, g, and Δ_{IS} (cm^{-1}), respectively, and are given in their unscaled values.] We may rationalize the misclassification of the 35 456.915 level by noting that in forming the training set one of the two existing $5f^7 6d 7p + 5f^8 6d$ levels was removed from the original even-parity data set leaving

$$35\,378.325,\quad 5.5,\quad 1.366,\quad -1.041$$

as the only $5f^7 6d 7p + 5f^8 6d$ level in the training set. We note that the E, J, g, and Δ_{IS} values of this level are very similar to those of the misclassified level at 35 456.915. Since the network acts as a nearest-neighbor classifier, a

TABLE I. Training percentages of counterpropagation neural networks for even- and odd-parity four feature energy levels of Cm II. Numbers in the table are the percentage of correct predictions when levels of a data set are recalled through a trained network. Numbers in parentheses are the number of misclassified levels.

Method	Even levels: All levels (74 levels)	Even levels: Training set[a] (59 levels)	Even levels: Test set[b] (15 levels)	Odd levels: All levels (82 levels)	Odd levels: Training set[a] (65 levels)	Odd levels: Test set[b] (17 levels)
Uniflow	100.0 (0)	96.6 (2)	80.0 (3)	97.5 (2)	100.0 (0)	76.4 (4)
Category learning	100.0 (0)	96.6 (2)	93.3 (1)	100.0 (0)	98.4 (1)	82.3 (3)

[a]Training set denotes the subset of classified energy levels used to train a network which is then used to predict configurations of levels in the test set.

[b]Test set denotes the subset of classified energy levels which were treated as unknowns and recalled through a network trained with the training set.

misclassification results. The above observations illustrate the importance of having more than one level belonging to a configuration in a training set. It is interesting to note that the configuration of the $5f^76d7p+5f^86d$ level which was removed from the training set (and therefore placed in the test set) was also incorrectly predicted. In regards to the misclassified level at 35 556.240 we note that its Δ_{IS} is a borderline value between those expected for $5f^76d7p$ and $5f^77s7p$ configurations (see Fig. 5 of Ref. [7]). This by itself cannot explain the error since isotope shifts overlap for the $5f^76d7p$, $5f^87s$, and $5f^77s7p$ configurations and there are levels belonging to these configurations which are classified correctly by the network. Nonetheless, it may serve as a useful rationalization.

B. Uniflow networks; even-parity levels

We next consider the performance of uniflow networks on even-parity levels. From Table I we see that two levels are misclassified in the training set. In fact, these two levels,

35 556.240, 3.5, 1.858, −0.743

$5f^76d7p$

and

40 966.185, 3.5, 1.778, −0.743

$5f^77s7p$

both give the same ambiguous output (i.e., a definitive category prediction cannot be made) of the type discussed in Ref [1]. This is somewhat surprising since there are no ambiguous classifications when *all* even-parity levels are used to train a uniflow network. This would tend to rule out improper data scaling as a cause of the ambiguity. In order to investigate this we tried several scalings, all of which were similar to but slightly different from (for example, dividing g by 9 and 11) the scaling above. All scalings gave the same results. We therefore conclude—as mentioned at the end of Sec. I—that this is a manifestation of PE's in the competitive layer taking responsibility for two or more training inputs which belong to different categories. We note in passing that the 35 556.240 level was also misclassified by the category learning network and that perhaps its Δ_{IS} is playing a role in this. In the test set the levels

35 280.665, 2.5, 2.255, −0.762

$5f^76d7p$ classified as $5f^87s$

and

35 778.500, 3.5, 1.130, −1.106

$5f^76d7p+5f^86d$ classified as $5f^86d$

are misclassified and the level

41 130.390, 4.5, 1.70, −0.698

$5f^77s7p$

gives ambiguous output. The first of these may be rationalized by noting the large overlap in isotope-shift values between the $5f^76d7p$ and $5f^87s$ configurations [7]; see the above discussion. The second level has been discussed above. The third level seems unremarkable, although we note that its E, J, g, and Δ_{IS} values are quite similar to the two ambiguous levels at 35 556.240 and 40 966.185.

C. Category-learning networks; odd-parity levels

The training-set level

28 079.390, 5.5, 1.32, −0.757 ,

$5f^76d7s$ classified as $5f^87p$,

is misclassified. The isotope-shift value of −0.757, however, does *not* fall in the range of Δ_{IS} values expected for the $5f^76d7s$ configuration [7] but does fall in the range expected for $5f^87p$. It therefore appears that the network has learned the correct isotope-shift "rule" in classifying the level as $5f^87p$. In the test set the category-learning network misclassifies three levels:

30 550.820, 1.5, 1.045, −0.955

$5f^76d^2$ classified as $5f^87p$

and

33 802.330, 4.5, 1.38, −0.925

$5f^87p$ classified as $5f^76d^2$

and

37 528.470, 0.5, −0.28, −0.854

$5f^87p$ classified as $5f^76d^2$.

The last two of these were misclassified by the uniflow network, and the isotope shifts of all three are in a range of overlap for the $5f^76d^2$ and $5f^87p$ configurations.

D. Uniflow networks; odd-parity levels

The levels

24 078.900, 5.5, 1.72, −0.978

$5f^76d^2$

and

27 065.085, 5.5, 1.51, −0.972

$5f^87p$

give rise to ambiguous output from a uniflow network trained with all classified levels. Both levels are also in the training set where they do *not* give rise to ambiguous outputs. As in the even-parity case we tried several scalings, all of which gave the same results. We therefore conclude that competitive layer PE's are behaving differently in the two networks; apparently in the all-level network some are taking responsibility for inputs belonging to different categories. The training-set uniflow network also results in ambiguous output for the levels

31 947.675, 3.5, 1.478, −0.859

$5f^87p$

and

35 291.395, 3.5, 1.382, −0.774

$5f^87p$

of the test set. The other two test-set levels which are misclassified are

33 802.330, 4.5, 1.38, −0.925

$5f^87p$ classified as $5f^76d^2$

and

37 528.470, 0.5, −0.28, −0.854

$5f^87p$ classified as $5f^76d^2$.

Both of these levels belonging to configurations for which there is a large overlap of isotope-shift values [7]. In regards to the second of these levels we note that there is only one level in the training set which has a negative g value and that it belongs to the $5f^76d^2$ configuration.

E. Configuration predictions of unclassified levels

Finally, we turn to the classification of unknown Cm II energy levels. The classifications were obtained by recalling unknown levels through networks which were trained on all classified levels of a given parity. The complete list of assignments can be obtained from the author on request—here we summarize the results. With the exception of six odd-parity levels, the uniflow and category-learning networks give the same predictions. Each of these six levels was classified as $5f^76d7s$ by the uniflow network and as $5f^87p$ by the category learning network. Of these six levels those at 34 023.550, 39 466.380, 40 285.940, and 41 134.885 all have isotope shifts consistent with the $5f^76d7s$ configuration. The level at 35 397.720 has a $\Delta_{\rm IS}$ consistent with the $5f^87p$ configuration, while the level at 38 472.695 has a $\Delta_{\rm IS}$ in between the ranges expected for the two configurations.

There is no angular-momentum coupling information (i.e., L, S, or J) given for any of the unknowns [7] and therefore such information cannot be used to check the configuration predictions for consistency. All predicted configurations of even-parity unknowns are consistent with Table IV of Ref. [7] which gives the lowest energy level of each configuration. For example, the unknown at 36 695.555, which is predicted as $5f^86d$, is higher in energy than the 17 150.790 value given for the lowest level of the $5f^86d$ configuration. The odd-parity unknown at 26 525.110 is lower in energy than the 27 065.085 value given for the lowest level of the $5f^87p$ configuration. It is likely that this level has been misclassified since its $\Delta_{\rm IS}$ is consistent with both the $5f^87p$ and $5f^76d^2$ configurations. Finally, all of the odd-parity configurations are subject to the following caveat. Levels belonging to all configurations expected below 50 000 have been identified with the exception of the odd-parity $5f^9$ configuration which is predicted to start around 26 000 [7]. Since all of the odd-parity unknowns in our study lie between 26 525.11 and 48 001.690 any of them could belong to the $5f^9$ configuration.

In order to obtain quantitative estimates of confidence levels for the predicted configurations of the unclassified levels, we performed the following set of computer experiments [9]. For the even-parity levels nine disjoint test sets comprised of eight randomly chosen energy levels were formed. The 66 energy levels remaining after each test-set selection served as the corresponding training set. In this way all but two energy levels in the total set of 74 served as a test-set level for one of the computer experi-

TABLE II. Values of performance measures averaged over test and training sets for Cm II even-parity energy levels. The first number of each pair is the value for the uniflow network and the second is that for the category learning network.

Category	Sensitivity	Specificity	PPV	FAR
		(a) Test sets		
f^87s	0.95, 1.00	0.98, 1.00	1.00, 1.00	0.02, 0.00
f^86d	1.00, 0.95	1.00, 0.98	0.95, 0.94	0.00, 0.02
f^77s7p	0.83, 0.83	0.97, 0.97	0.91, 0.91	0.03, 0.03
f^76d7p	0.72, 0.84	0.91, 0.94	0.81, 0.80	0.09, 0.06
$f^76d7p+f^86d$	0.00, 0.00	0.97, 0.97	0.00, 0.00	0.03, 0.03
all levels	0.86, 0.89	0.99, 0.97	0.91, 0.90	0.01, 0.03
		(b) Training sets		
f^87s	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
f^86d	1.00, 0.99	1.00, 0.998	1.00, 1.00	0.00, 0.002
f^77s7p	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
f^76d7p	1.00, 1.00	1.00, 1.00	1.00, 0.99	0.00, 0.00
$f^76d7p+f^86d$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
all levels	1.00, 0.998	1.00, 0.9996	1.00, 0.998	0.00, 0.0004

TABLE III. Values of performance measures averaged over test and training sets for Cm II odd-parity energy levels. The first number of each pair is the value for the uniflow network and the second is that for the category-learning network.

Category	Sensitivity	Specificity	PPV	FAR
		(a) Test sets		
$f^7 7s^2$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
$f^7 6d7s$	0.96, 0.96	0.94, 0.94	0.93, 0.96	0.06, 0.06
$f^7 6d^2$	0.71, 0.71	0.92, 0.92	0.71, 0.71	0.08, 0.08
$f^8 7p$	0.44, 0.56	0.87, 0.89	0.70, 0.60	0.13, 0.11
all levels	0.80, 0.82	0.94, 0.94	0.86, 0.84	0.06, 0.06
		(b) Training sets		
$f^7 7s^2$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
$f^7 6d7s$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
$f^7 6d^2$	0.94, 1.00	0.98, 1.00	1.00, 1.00	0.02, 0.00
$f^8 7p$	0.94, 1.00	0.99, 1.00	1.00, 1.00	0.01, 0.00
all levels	0.98, 1.00	0.99, 1.00	1.00, 1.00	0.01, 0.00

ments. A given network was trained and tested on each training-testing set pair. The average performance over the nine test sets provides a measure of the probability of correct classification [10]. For the odd-parity levels nine disjoint test sets comprised of nine randomly chosen energy levels were formed. In this way all but one energy level in the total set of 82 served as a test-set level.

In addition to calculating the probability of correct classification for each configuration we have calculated several other performance measures which we now briefly describe [11]. For a given category, four possible alternatives exist for the predicted configuration of a level. The first alternative is a true positive prediction ($P_{\rm TP}$), in which the network correctly predicts that the level belongs to a given category. The second is a false positive prediction ($P_{\rm FP}$), in which the network incorrectly predicts that the level belongs to the given category. The third possibility is a false negative prediction ($P_{\rm FN}$), in which the network incorrectly predicts that the level belongs to a different category. The fourth possibility is a true negative prediction ($P_{\rm TN}$), in which the network correctly predicts that the level belongs to a different category. The sensitivity (also called the true positive ratio or the recall) is given by $P_{\rm TP}/(P_{\rm TP}+P_{\rm FN})$; when multiplied by 100 it is the percent of correct predictions. The specificity is given by $P_{\rm TN}/(P_{\rm TN}+P_{\rm FN})$ and is also called the true negative ratio. The positive predictive value (PPV), or precision, is $P_{\rm TP}/(P_{\rm TP}+P_{\rm FP})$. Finally, the false alarm rate (FAR), or false positive ratio is $(1-S)=P_{\rm FP}/(P_{\rm FP}+P_{\rm TN})$, where S is the specificity.

The sensitivity, specificity, PPV, and FAR were calculated for each configuration of each network and then averaged over the networks. The results are presented in Tables II and III and give estimates of the reliability of the predicted configurations of the unclassified energy levels. With the exception of the even-parity $5f^7 6d7p+5f^8 6d$ configuration and the odd-parity $5f^8 7p$ configuration, the test-set sensitivities are all 0.71 or greater and the corresponding PPV's are generally larger than the sensitivities. While the sensitivities of the two exceptional configurations are quite low, their FAR's are very small. This indicates that although relatively few of the unclassified levels which actually belong to the $5f^7 6d7p+5f^8 6d$ and $5f^8 7p$ configurations will be classified correctly, an even fewer number of unclassified levels which do not actually belong to these configurations will be predicted as belonging to them. None of the unclassified even-parity levels have been predicted to belong to the $5f^7 6d7p+5f^8 6d$ configuration. Sixty-eight unclassified odd-parity levels have been predicted to belong to the $5f^8 7p$ configuration. Finally, it is encouraging that the largest FAR for all Cm II configurations is only 0.13.

III. APPLICATION TO Pu I ENERGY LEVELS

In this section we discuss the application of uniflow and category learning networks to the classification of Pu I energy levels. Data are taken from Blaise, Fred, and Gutmacher [12]. Considering only four feature levels leaves us with an odd-parity data set consisting of 288 levels representing six configurations ($5f^5 6d7s^2$, $5f^5 6d^2 7s$, $5f^6 7s7p$, $5f^7 7s$, $5f^6 6d7p$, and $5f^5 6d7s8s$) and an even-parity data set consisting of 221 levels representing eight configurations ($5f^6 7s^2$, $5f^6 6d7s$, $5f^5 7s^2 7p$, $5f^5 6d7s7p$, $5f^6 7s8s$, $5f^6 6d^2$, $5f^4 6d^2 7s^2$, and $5f^5 6d^2 7p$). The networks constructed for analyzing the odd-parity data set consisted of an input layer, a normalizing layer, a competitive layer, and an output layer having four, five, 288, and six PE's, respectively. Similarly, networks for the even-parity data have layers consisting of four, five, 221, and eight PE's. The networks were run on the same computer as were the Cm II networks, utilizing the same software package, number of passes through data sets, and data scaling.

The training results of various networks are summarized in Table IV. The formation and use of training and test sets is completely analogous to that described in Sec. II for Cm II. Actual numbers of PE's in competitive layers can be obtained from the parenthetical entries in Table IV. We now make some observations which rationalize some of the mistakes reflected in Table IV.

TABLE IV. Training percentages of counterpropagation neural networks for even- and odd-parity four feature energy levels of Pu I. Numbers in the table are the percentage of correct predictions when levels of a data set are recalled through a trained network. Numbers in parentheses are the number of misclassified levels.

Method	Even levels: All levels (221 levels)	Even levels: Training set[a] (176 levels)	Even levels: Test set[b] (45 levels)	Odd levels: All levels (288 levels)	Odd levels: Training set[a] (230 levels)	Odd levels: Test set[b] (58 levels)
Uniflow	98.6 (3)	98.3 (3)	73.3 (12)	98.6 (4)	98.2 (4)	77.6 (13)
Category learning	90.5 (21)	91.4 (15)	77.7 (10)	91.6 (24)	92.6 (17)	75.8 (14)

[a]Training set denotes the subset of classified energy levels used to train a network which is then used to predict configurations of levels in the test set.
[b]Test set denotes the subset of classified energy levels which were treated as unknowns and recalled through a network trained with the training set.

A. Category-learning networks; even-parity levels

The category-learning network misclassifies 21 levels in the even-parity all-level data set. These levels with their J, g, and $\Delta_{\rm IS}$ values, and actual and predicted configurations are given in Table V. The first observation concerning these levels is that they are at relatively high energy; the lowest is at 25 655.090. As energy increases, configuration mixing increases, g values tend toward unity, and $\Delta_{\rm IS}$ values tend toward equality [12]. Thus one would expect higher energy levels to be generally more difficult to classify. In regards to the levels in part (a) of the table, we note that levels belonging to the $5f^67s^2$ configuration, if pure, should have a $\Delta_{\rm IS}$ of 480, while those belonging to the $5f^66d7s7p$ configuration, if pure, should have a $\Delta_{\rm IS}$ of 430 [12]. These two $\Delta_{\rm IS}$ values are more similar than for any other pair of even-parity configurations. Thus it may not be surprising that the network has difficulty discriminating between these two configurations. The same two configurations are involved in part (b) of the table. For this level the $\Delta_{\rm IS}$ is closer to that of the predicted configuration. As regards part (c) of the table, we note that pure $5f^57s^27p$ levels should have a $\Delta_{\rm IS}$ of 725 and that pure $5f^67s^2$ levels should have a $\Delta_{\rm IS}$ of 480 [12]. All four levels have $\Delta_{\rm IS}$ values closer to the value of 480 of the predicted $5f^67s^2$ configuration than to 725. A similar argument holds for the first level of part (d) and for part (g) of the table; these levels are denoted with a footnote indicator in Table V and in all subsequent tables. (See the above discussion for $\Delta_{\rm IS}$'s of the relevant configurations; the $\Delta_{\rm IS}$ of pure $5f^66d7s$ levels is 250 [12].) Unfortunately, this isotope-shift rationalization argument does not hold for the three levels in parts (e) and (f) of the table. (The $\Delta_{\rm IS}$ of pure $5f^67s8s$ levels is 360 [12].) Using $\Delta_{\rm IS}$ values based on pure configurations may not be the most enlightening way of rationalizing our results since there is a large amount of configuration mixing in many of the Pu I levels. Although we will continue to point out when this rationalization is valid, we will also consider an argument based on the observed ranges of $\Delta_{\rm IS}$ values of the various configurations of Pu I. Using data from Ref. [12], we have compiled the experimentally observed ranges of $\Delta_{\rm IS}$ values which we give here for future reference (numbers in parentheses are the pure-configuration $\Delta_{\rm IS}$ values given in Ref. [12]):

TABLE V. Even-parity levels of Pu I misclassified by a category learning network trained with all even-parity levels

E(cm^{-1})	J	g	$\Delta_{\rm IS}$(10^{-3} cm^{-1})
(a) $5f^56d7s7p$ classified as $5f^67s^2$			
28 295.380[a,b]	4	1.155	475
30 083.102[b]	5	1.15	446
30 461.399[a,b]	5	0.99	461
30 544.187[a,b]	3	0.96	502
31 628.619[a,b]	6	1.11	509
31 732.582[b]	5	1.049	432
32 404.416[a,b]	5	1.005	499
32 440.827[b]	6	1.12	420
34 045.560[a,b]	7	1.17	503
35 323.031[b]	6	1.09	439
(b) $5f^67s^2$ classified as $5f^56d7s7p$			
33 884.230[a]	7	1.105	340
(c) $5f^57s^27p$ classified as $5f^67s^2$			
30 113.280[a,b]	6	1.05	541
31 413.230[a,b]	3	0.742	526
31 881.871[a,b]	7	1.12	511
36 128.425[a,b,]	7	1.06	547
(d) $5f^57s^27p$ classified as $5f^56d7s7p$			
31 810.821[a,b,]	2	0.865	533
36 230.135	9	1.255	580
(e) $5f^66d7s$ classified as $5f^67s^2$			
25 655.090[b]	4	1.165	366
30 016.377	2	1.14	333
(f) $5f^56d7s7p$ classified as $5f^67s8s$			
33 180.043[b]	2	1.79	445
(g) $5f^66d7s$ classified as $5f^56d7s7p$			
25 979.424[a]	1	1.26	384

[a]Denotes a level for which the isotope-shift rationalization (based on pure configurations) discussed in the text is valid.
[b]Denotes a level for which the isotope-shift rationalization (based on experimentally observed ranges of $\Delta_{\rm IS}$ values within configurations) discussed in the text is valid.

Even parity		
$5f^67s^2$	340–566	(480)
$5f^66d7s$	237–384	(250)
$5f^57s^27p$	507–698	(725)
$5f^56d7s7sp$	414–546	(430)
$5f^67s8s$	333–446	(360)
$5f^66d^2$	115–349	(100)
$5f^46d^27s^2$	535	(840)
$5f^56d^27p$	390–403	(160)
Odd parity		
$5f^56d7s^2$	420–683	(665)
$5f^56d^27s$	370–546	(410)
$5f^77s$	273	(187)
$5f^66d7p$	293–383	(100)
$5f^56d7s$	503	(500)

There is only one observed classified level for each of the $5f^46d^27s^2$, $5f^77s$, and $5f^56d7s$ configurations and hence, no range of Δ_{IS} values can be given for them. Using the above ranges, we see that for all but four levels in Table V (33 884.230, 36 230.135, 30 016.377, and 25 979.424), the Δ_{IS} is consistent with the observed range of values for the predicted configuration. (Levels for which this is the case are denoted by a [b] in this and all subsequent tables.) We thus have a situation very similar to that which occurred for Cm II. It is perhaps not surprising that the network cannot discriminate between the proper configurations of these levels.

Table VI shows the even-parity training-set levels that were misclassified by the category-learning network. Levels marked with an asterisk have already been discussed in relation to Table V. This leaves parts (b), (d), (f), (h), and (j) remaining to be discussed. The isotope-shift argument based on pure configurations fails for all six of these levels. Unfortunately, the isotope shifts of only two of these levels are consistent with the observed ranges of Δ_{IS} values for the predicted configurations. With regards to parts (h) and (i) of the table, we note that the energy values of these two levels are inconsistent with the lowest observed energies of their predicted configurations. That is, the levels at 31 471.542 and 24 016.378 are both lower in energy than the energy (31 572 [12]) of the lowest level of the predicted $5f^67s8s$ configuration; the network's prediction is nonsensical.

Table VII shows the even-parity test-set levels that were misclassified by the category learning network. Levels marked with an asterisk have already been discussed in relation to Table V. We make the following observations concerning the various parts of the table.

(a) The Δ_{IS} is consistent with the actual configuration, not the predicted configuration. There is one $5f^67s^2$ level in the training set (at 9 772.532) with $J=g=0$. However, there are two $5f^56d7s7p$ levels in the training set (at 28 763.085 and 30 649.882) with $J=g=0$. Since these two levels are much closer in energy to that of the level in question, a misclassification results.

(c) The Δ_{IS} of 533 is much closer to that expected for the predicted $5f^66d7s7p$ configuration.

(d) The Δ_{IS} of 253 is almost that expected for the actual configuration; neither of the Δ_{IS} rationalizations (based on pure configurations or on observed ranges) invoked above works. In addition, the energy of this level is lower than the energy of the lowest level of the predicted configuration [12]; this is a nonsensical result.

(e) See paragraph (d).

(f) See paragraph (a).

(g) The Δ_{IS} is consistent with the actual configuration, not the predicted configuration. There are no $5f^67s8s$ levels in the training set with $J=g=0$. There is one $5f^66d^2$ level in the training set (at 34 006.573) with $J=g=0$; this level has a Δ_{IS} close to the Δ_{IS} of the level

TABLE VI. Even-parity training-set levels of Pu I misclassified by a category-learning network. Levels marked with an asterisk are found in Table V; those with a dagger are also found in Table V, but with a different predicted configuration; and levels marked with a double dagger are lower in energy than the lowest level of the predicted configuration.

E (cm^{-1})	J	g	Δ_{IS} (10^{-3} cm^{-1})
(a) $5f^56d7s7p$ classified as $5f^67s^2$			
30 083.102[b,*]	5	1.15	446
31 628.619[a,b,*]	6	1.11	509
31 732.582[b,*]	5	1.049	432
34 045.560[a,b,*]	7	1.17	503
(b) $5f^67s^2$ classified as $5f^56d7s7p$			
22 339.429[b]	2	1.049	481
(c) $5f^57s^27p$ classified as $5f^67s^2$			
31 881.871[a,b,*]	7	1.12	511
(d) $5f^67s^2$ classified as $5f^57s^27p$			
24 753.684	4	0.975	481
24 921.671[b]	5	1.034	555
(e) $5f^57s^27p$ classified as $5f^56d7s7p$			
36 128.425[a,†]	7	1.06	547
36 230.135*	9	1.255	580
(f) $5f^56d7s7p$ classified as $5f^57s^27p$			
29 976.039	4	1.07	489
(g) $5f^66d7s$ classified as $5f^67s^2$			
30 016.377*	2	1.14	333
(h) $5f^56d7s7p$ classified as $5f^67s8s$			
31 471.542[‡]	1	2.188	546
(i) $5f^66d7s$ classified as $5f^56d7s7p$			
25 979.424[a,*]	1	1.26	384
(j) $5f^66d7s$ classified as $5f^67s8s$			
24 016.378[‡]	5	1.56	274

[a]Denotes a level for which the isotope-shift rationalization (based on pure configurations) discussed in the text is valid.
[b]Denotes a level for which the isotope-shift rationalization (based on experimentally observed ranges of Δ_{IS} values within configurations) discussed in the text is valid.

in question and a misclassification results.

(i) The Δ_{IS} values of both levels are consistent with the actual configuration, not the predicted configuration. In regards to the level at 17 336.413, we note that there are no $5f^6 6d7s$ levels in the training set with $J = g = 0$. As noted in paragraph (g), there is one $5f^6 6d^2$ level in the training set with $J = g = 0$ and a misclassification results. Both levels are lower in energy than the lowest level of the predicted configuration [12].

B. Uniflow networks; even-parity levels

We turn next to the performance of uniflow networks with even-parity levels. In the all-level network three levels give ambiguous output:

19 337.431, 1, 2.41, 256

$5f^6 6d7s$

and

31 471.542, 1, 2.188, 546

$5f^6 6d7s7p$

and

31 572.610, 1, 2.403, 446

$5f^6 7s8s$.

These three levels also give rise to ambiguous output from a uniflow network trained with the training set. The distinguishing features of these levels are their relatively large g values and their common J value of 1. As with Cm II, we tried several different scalings without success. We note, however, that all three levels are classified correctly with the category-learning network.

Table VIII shows the even-parity test-set levels that were misclassified by the uniflow network. Levels marked with an asterisk give rise to predicted configurations identical to those obtained from the corresponding category-learning network. The level at

TABLE VII. Even-parity test-set levels of Pu I misclassified by a category-learning network. Levels marked with an asterisk are found in Table V; those with a dagger are found in Table V, but with a different predicted configuration; and levels marked with a double dagger are lower in energy than the lowest level of the predicted configuration.

E (cm^{-1})	J	g	Δ_{IS} (10^{-3} cm^{-1})
(a) $5f^6 7s^2$ classified as $5f^5 6d7s7p$			
32 324.169[b]	0	0.000	480
(b) $5f^5 7s^2 7p$ classified as $5f^6 7s^2$			
30 113.280[a,b,*]	6	1.05	541
(c) $5f^5 7s^2 7p$ classified as $5f^5 6d7s7p$			
23 806.381[a,b]	1	0.094	533
(d) $5f^6 6d7s$ classified as $5f^5 6d7s7p$			
14 341.947[‡]	2	0.852	253
(e) $5f^6 6d7s$ classified as $5f^6 7s8s$			
26 476.068[‡]	4	1.605	260
(f) $5f^6 7s8s$ classified as $5f^5 6d7s7p$			
35 032.090	0	0.000	391
(g) $5f^6 7s8s$ classified as $5f^6 6d^2$			
33 304.400[b]	0	0.000	344
(h) $5f^6 7s^2$ classified as $5f^6 6d7s$			
33 884.230[a,b,†]	7	1.105	340
(i) $5f^6 6d7s$ classified as $5f^6 6d^2$			
17 336.413[b,‡]	0	0.000	257
25 293.751[b,‡]	3	0.965	278

[a]Denotes a level for which the isotope-shift rationalization (based on pure configurations) discussed in the text is valid.
[b]Denotes a level for which the isotope-shift rationalization (based on experimentally observed ranges of Δ_{IS} values within configurations) discussed in the text is valid.

TABLE VIII. Even-parity test-set levels of Pu I misclassified by a uniflow network. Levels marked with an asterisk are found in Table VII; those with a dagger are also found in Table VII, but with a different predicted configuration; and levels with a double dagger are lower in energy than the lowest level of the predicted configuration.

E (cm^{-1})	J	g	Δ_{IS} (10^{-3} cm^{-1})
(a) $5f^6 7s^2$ classified as $5f^5 6d7s7p$			
25 707.348[b]	2	0.72	516
27 805.163[b]	5	1.024	513
32 324.169[b,*]	0	0.000	480
33 884.230[a,†]	7	1.105	340
(b) $5f^5 7s^2 7p$ classified as $5f^6 7s^2$			
30 113.280[a,b,*]	6	1.05	541
(c) $5f^5 7s^2 7p$ classified as $5f^5 6d7s7p$			
23 806.381[a,b,*]	1	0.094	533
(d) $5f^6 6d7s$ classified as $5f^5 6d7s7p$			
14 341.947[*,‡]	2	0.852	253
(e) $5f^6 7s8s$ classified as $5f^5 6d7s7p$			
35 032.090*	0	0.000	391
(f) $5f^6 7s8s$ classified as $5f^6 6d^2$			
33 304.400[b,*]	0	0.000	344
(g) $5f^6 6d7s$ classified as $5f^6 6d^2$			
17 336.413[b,*,‡]	0	0.000	257
28 793.800[‡]	3	1.083	366
(h) $5f^5 6d7s7p$ classified as $5f^6 7s^2$			
26 205.589[a,b]	3	0.701	462

[a]Denotes a level for which the isotope-shift rationalization (based on pure configurations) discussed in the text is valid.
[b]Denotes a level for which the isotope-shift rationalization (based on experimentally observed ranges of Δ_{IS} values within configurations) discussed in the text is valid.

33 884.230 (marked with a dagger) is classified as $5f^56d7s7p$ by the uniflow network but as $5f^66d7s$ by the category-learning network. There are four levels misclassified by the uniflow network that were classified correctly by the category-learning network. These levels are noted by the absence of an asterisk. Among these levels the Δ_{IS} rationalization argument based on pure configurations holds for the level at 26 205.589 but fails for those at 25 707.348, 27 805.163, and 28 793.800. The Δ_{IS} rationalization based on observed ranges works better, being valid for the first three of these levels, but failing for the one at 28 793.800. The levels at 14 341.947, 17 336.413, and 28 793.800 are all inconsistent with the lowest levels of their predicted configurations [12].

TABLE IX. Odd-parity levels of Pu I misclassified by a category-learning network trained with all odd-parity levels. Levels marked with a double dagger are lower in energy than the lowest level of the predicted configuration.

E (cm^{-1})	J	g	Δ_{IS} (10^{-3} cm^{-1})
(a) $5f^56d7s^2$ classified as $5f^56d^27s$			
19 426.512	3	1.435	550
24 437.792	4	1.50	584
26 575.338[b]	7	1.15	538
26 844.163[a,b]	4	1.020	528
28 385.761[a,b]	6	1.025	514
28 890.990[a,b]	7	1.272	470
(b) $5f^56d7s^2$ classified as $5f^67s7p$			
24 644.996	3	1.195	552
28 749.920[b]	2	1.48	504
30 932.959[a,b]	6	1.34	420
(c) $5f^56d^27s$ classified as $5f^67s7p$			
22 518.312[b]	2	1.35	419
25 113.744[b]	6	1.302	438
25 121.896[b]	1	1.444	428
25 397.206[b]	1	0.776	453
25 660.792[b]	1	1.146	452
26 317.729[b]	4	1.18	405
27 651.193[b]	2	1.570	430
27 909.524[b]	4	1.27	427
28 021.637[b]	5	1.40	407
(d) $5f^67s7p$ classified as $5f^56d^27s$			
27 334.422[a,b]	7	1.24	417
28 595.088[a,b]	6	1.20	380
31 130.605[a,b]	1	2.342	397
31 151.870	2	1.632	350
(e) $5f^67s7p$ classified as $5f^66d7p$			
30 319.724[b,‡]	3	1.010	371
31 233.379[‡]	4	1.09	399

[a]Denotes a level for which the isotope-shift rationalization (based on pure configurations) discussed in the text is valid.
[b]Denotes a level for which the isotope-shift rationalization (based on experimentally observed ranges of Δ_{IS} values within configurations) discussed in the text is valid.

C. Category-learning networks; odd-parity levels

Table IX shows the 24 odd-parity levels that are misclassified by the category-learning network trained with all odd-parity levels. The Δ_{IS} rationalization argument based on pure configurations which is valid for many even-parity levels unfortunately does not work as often for the odd-parity levels (the argument holds for seven of the 24 levels). However, the argument based on observed ranges is valid for 19 levels. The levels at 30 319.724 and 31 233.379 are both lower in energy than the lowest level of their predicted $5f^66d7p$ configuration [12].

Table X shows the odd-parity training-set levels that are misclassified by the category-learning network. The Δ_{IS} rationalization argument based on pure configurations is valid for only three of the 17 levels, while the argument based on observed ranges is valid for 13 of the levels. There are not as many different types of misclassifications in the training set as in the all-level set. There are no $5f^56d7s^2$ levels classified as $5f^67s7p$ and there are no $5f^67s7p$ levels classified as $5f^66d7p$. The two $5f^56d7s^2$ levels at 24 644.996 and 28 749.920 which

TABLE X. Odd-parity training-set levels of Pu I misclassified by a category-learning network. Levels marked with an asterisk are found in Table IX; those with a dagger are also found in Table IX, but with a different predicted configuration; and levels marked with a double dagger are lower in energy than the lowest level of the predicted configuration.

E (cm^{-1})	J	g	Δ_{IS} (10^{-3} cm^{-1})
(a) $5f^56d7s^2$ classified as $5f^56d^27s$			
19 426.512*	3	1.435	550
24 437.792*	4	1.50	584
24 644.996*,†	3	1.195	552
26 575.448[b,*]	7	1.15	538
28 749.920[a,b,*,†]	2	1.48	504
(b) $5f^56d^27s$ classified as $5f^67s7p$			
22 518.312[b,*]	2	1.35	419
23 281.721[b]	5	1.235	452
25 100.598[b]	2	1.43	408
25 397.206[b,*]	1	0.776	453
25 660.792[b,*]	1	1.146	452
25 959.849[b]	1	1.037	441
26 149.538[b]	4	1.36	424
26 317.729[b,*]	4	1.18	405
27 651.193[b,*]	2	1.57	430
(c) $5f^67s7p$ classified as $5f^56d^27s$			
22 429.984[a,b]	4	1.279	433
23 274.858[a,b]	4	1.604	418
25 074.585	4	1.507	324

[a]Denotes a level for which the isotope-shift rationalization (based on pure configurations) discussed in the text is valid.
[b]Denotes a level for which the isotope-shift rationalization (based on experimentally observed ranges of Δ_{IS} values within configurations) discussed in the text is valid.

are classified as $5f^67s7p$ in the all-level case are classified as $5f^56d^27s$ in the training-set case. The reason for this is not readily apparent. The $5f^56d7s^2$ level at 30 932.959 is not present in the training set and is a member of the test set. The two $5f^67s7p$ levels at 30 319.724 and 31 233.379 are classified correctly in the training set.

Table XI shows the odd-parity test-set levels that are misclassified by the category-learning network. As in the even-parity test-set case, there are very few misclassified levels which are also misclassified in the all-level case; only three such levels exist. The Δ_{IS} rationalization argument based on pure configurations works considerably better here than for previous odd-parity cases, being valid for one-half of the levels. The rationalization based on observed ranges holds for all but two levels. As with the even-parity case it is possible to find arguments which rationalize some, although not all, of the misclassifications. For example, if we consider the level at 28 385.761, we note that in the training set the highest energy level with $J=6$ belonging to the actual $5f^56d7s^2$ configuration is at 26 443.391 (with $\Delta_{IS}=617$ and $g=1.045$), while there is a $J=6$ level of the predicted $5f^56d^27s$ configuration at 28 036.676 ($\Delta_{IS}=475$, $g=1.216$). Taken together, it appears that the four values E, J, g, and Δ_{IS} for the 28 385.761 level are more similar to those of the 26 443.391 level—hence the misclassification. An entirely analogous argument holds for the level at 30 932.959. The situation for the two other $J=6$ levels in the table is not as clear-cut and no definitive rationalization can be made for these levels. In regards to the remaining levels in the table, it is sometimes possible to find training-set levels belonging to the predicted configuration that appear to be similar to the level in question. However, finding the most similar level by inspection is difficult at best. We will return to this point in the conclusion of this paper.

D. Uniflow networks; odd-parity levels

In the all-level network four levels give ambiguous output: 17 500.977, 1, 2.258, 523

$5f^56d7s^2$

and

21 307.390, 1, 2.360, 405

$5f^56d^27s$

and

30 929.516, 1, 2.26, 391

$5f^67s7p$

TABLE XI. Odd-parity test-set levels of Pu I misclassified by a category-learning network. Levels marked with an asterisk are found in Table IX.

E (cm^{-1})	J	g	Δ_{IS} (10^{-3} cm^{-1})
(a) $5f^56d7s^2$ classified as $5f^56d^27s$			
24 012.505	6	1.248	574
25 839.917[b]	6	1.25	537
28 385.761[a,b,*]	6	1.025	514
(b) $5f^56d7s^2$ classified as $5f^67s7p$			
19 281.917	2	1.822	629
30 932.959[a,b,*]	6	1.34	420
(c) $5f^56d^27s$ classified as $5f^67s7p$			
20 769.512[b]	2	1.07	370
24 188.639[b]	1	0.667	401
27 869.060[b]	7	1.25	421
28 021.637[b,*]	5	1.40	407
(d) $5f^67s7p$ classified as $5f^56d^27s$			
17 045.776[a,b]	1	1.474	385
18 578.669[a,b]	1	1.932	402
27 228.191[a,b]	1	1.767	382
28 906.355[a,b]	3	1.23	532
29 295.313[a,b]	4	1.27	480

[a]Denotes a level for which the isotope-shift rationalization (based on pure configurations) discussed in the text is valid.
[b]Denotes a level for which the isotope-shift rationalization (based on experimentally observed ranges of Δ_{IS} values within configurations) discussed in the text is valid.

TABLE XII. Odd-parity test-set levels of Pu I misclassified by a uniflow network. Levels marked with an asterisk are those found in Table XI; those with a dagger are found in Table XI, but with a different predicted configuration.

E (cm^{-1})	J	g	Δ_{IS} (10^{-3} cm^{-1})
(a) $5f^56d7s^2$ classified as $5f^56d^27s$			
25 839.917[a,b,*]	6	1.25	537
(b) $5f^56d7s^2$ classified as $5f^67s7p$			
19 281.917[*]	2	1.822	629
30 932.959[a,b,*]	6	1.34	420
(c) $5f^56d^27s$ classified as $5f^67s7p$			
16 532.104[b]	3	0.3	450
(d) $5f^56d^27s$ classified as $5f^67s7p$			
20 769.512[b,*]	2	1.07	370
24 188.639[b,*]	1	0.667	401
27 869.060[b,*]	7	1.25	421
28 021.637[b,*]	5	1.40	407
(e) $5f^67s7p$ classified as $5f^56d7s^2$			
28 906.355[a,b,†]	3	1.23	532
(f) $5f^67s7p$ classified as $5f^56d^27s$			
17 045.776[a,b,*]	1	1.474	385
21 031.258	2	1.455	342
27 228.191[a,b,*]	1	1.767	382
29 295.313[a,b,*]	4	1.27	480

[a]Denotes a level for which the isotope-shift rationalization (based on pure configurations) discussed in the text is valid.
[b]Denotes a level for which the isotope-shift rationalization (based on experimentally observed ranges Δ_{IS} values within configurations) discussed in the text is valid.

and

31 130.605, 1, 2.342, 397

$5f^67s7p$.

These four levels also give rise to ambiguous output from a uniflow network trained with the training set. As in the even-parity case the distinguishing features of these levels are their relatively large g values and their common J value of 1. We again tried several different scalings without success. It is interesting to note that the level

23 766.136, 1, 2.162, 310

$5f^67s7p$

which also has $J=1$ and a relatively large g value, is correctly classified. In addition, each of the four levels with the exception of that at 31 130.605 is correctly classified by the category-learning network.

Table XII shows the odd-parity test-set levels that were misclassified by the uniflow network. Levels marked with an asterisk give rise to predicted configurations identical to those obtained from the corresponding category-learning network. The level at 28 906.355 (marked with a dagger) is classified as $5f^56d7s^2$ by the uniflow network but as $5f^56d^27s$ by the category-learning network. There are two levels misclassified by the uniflow network which are classified correctly by the category-learning network. These levels are noted by the absence of an asterisk. The isotope shifts of six of the 13 levels are consistent with the pure level isotope shifts of the predicted configurations. The Δ_{IS} argument based on observed ranges is valid for all but two levels.

E. Configuration predictions of unclassified levels

Finally, we turn to the classification of unknown Pu I energy levels. The classifications are obtained in a manner completely analogous to that for Cm II. The complete list of assignments can be obtained from the author—here we summarize the results. Unlike the Cm II unknowns where category learning and uniflow networks gave the same predictions for all but six levels, there is considerably more disagreement between the predictions of both networks for Pu I unknowns; 97 of the 255 even-parity unknowns gave rise to two different predicted configurations while 56 of the 230 odd-parity unknowns did so. There are nine even-parity unknowns whose Δ_{IS} is inconsistent with the observed range of both predicted configurations, 51 whose Δ_{IS} is consistent with only one of the predicted configurations, and 37 whose Δ_{IS} is consistent with both of the predicted configurations. There are 28 odd-parity unknowns whose Δ_{IS} is consistent with the observed range of Δ_{IS} values of only one of the predicted configurations and 27 unknowns whose Δ_{IS} value is consistent with both predicted configurations. In light of the above observations concerning the training results of networks where we noted that misclassifications frequently arise when the Δ_{IS} value of a level is consistent with the observed ranges of two configurations, we may legitimately cast doubt on the validity of the classifications of the 37 even-parity and 27 odd-parity unknowns which satisfy this condition. Of course, an unknown which gives rise to the same predicted configuration from both the category-learning and uniflow networks is not necessarily a level for which a highly certain prediction has been made. This may be seen by comparing Table VII with Table VII (and Table XI with Table XII). From these two tables we see that of the 12 levels in the test set which were misclassified by a uniflow network, seven of them were also misclassified by the category-learning network, the same configuration being predicted by each type of network.

As with Cm II there is no angular-momentum coupling information given for any of the unknowns. Therefore such information cannot be used to check the predicted configurations for consistency. Not all of the predicted configurations are consistent with Table III of Ref. [12] which gives the lowest energy of each configuration. The lowest levels of the even-parity configurations $5f^66d^2$, $5f^46d^27s^2$, and $5f^56d^27p$ lie at 31 710, 36 050, and 37 415, respectively, while the lowest level of the odd-parity $5f^66d7s$ configuration lies at 33 070 [12]. Levels predicted as one of these configurations which lie below the corresponding lowest levels are therefore incorrectly classified. In many of the cases where two configurations are predicted for a level these considerations resolve the ambiguity. (For example, the even-parity level at 32 869.165 would be assigned to the $5f^56d7s7p$ configuration.) Unfortunately, in some cases we are left with no predicted configuration—for example, the even-parity level at 26 449.501.

In order to obtain quantitative estimates of confidence levels for the predicted configurations of the unclassified levels, we performed a set of experiments analogous to those done for Cm II. For the even-parity levels 10 disjoint test sets comprised of 22 randomly chosen energy levels were formed. The 199 energy levels remaining after each test-set selection served as the corresponding training set. In this way all but one energy level in the total set of 221 served as a test-set level for one of the experiments. For the odd-parity levels nine disjoint test sets comprised of 32 randomly chosen energy levels were formed. In this way each of the energy levels in the total set of 288 served as a test-set level.

The sensitivity, specificity, PPV, and FAR were calculated for each configuration and then averaged over the networks. The results are presented in Tables XIII and XIV. There is a large variation in test-set sensitivity values among configurations, ranging from 0.00 to 0.92. The even-parity configurations $5f^66d^2$ and $5f^46d^27s^2$ and the odd-parity configurations $5f^77s$ and $5f^56d7s8s$ have sensitivities of 0.00, indicating that counterpropagation neural networks are not effective in correctly classifying levels belonging to these configurations. It is, however, somewhat encouraging that the FAR values of these configurations—as for all other configurations—are very small. None of the odd-parity unclassified levels are predicted as belonging to the $5f^77s$ or $5f^56d7s8s$ configurations. Fifteen even-parity unclassified levels are predicted as belonging to the $5f^66d^2$ configuration by uniflow networks; two of these levels have energies lower

TABLE XIII. Values of performance measures averaged over test and training sets for Pu I even-parity energy levels. The first number of each pair is the value for the uniflow networks and the second is that for the category-learning networks.

Category	Sensitivity	Specificity	PPV	FAR
		(a) Test sets		
$f^6 7s^2$	0.38, 0.53	0.89, 0.90	0.45, 0.35	0.11, 0.10
$f^6 6d7s$	0.81, 0.75	0.94, 0.92	0.94, 0.90	0.06, 0.08
$f^5 7s^2 7p$	0.76, 0.63	0.93, 0.89	0.78, 0.76	0.07, 0.11
$f^5 6d7s7p$	0.65, 0.61	0.85, 0.87	0.57, 0.70	0.15, 0.13
$f^6 7s8s$	0.60, 0.70	0.98, 0.99	0.55, 0.47	0.02, 0.01
$f^6 6d^2$	0.00, 0.00	0.99, 0.99	0.00, 0.00	0.01, 0.01
$f^4 6d^2 7s^2$	0.00, 0.00	0.99, 0.99	0.00, 0.00	0.01, 0.01
$f^5 6d^2 7p$	0.00, 1.00	0.99, 1.00	0.00, 1.00	0.01, 0.00
all levels	0.65, 0.64	0.85, 0.95	0.68, 0.64	0.05, 0.05
		(b) Training sets		
$f^6 7s^2$	1.00, 0.96	1.00, 0.99	1.00, 0.70	0.00, 0.01
$f^6 6d7s$	0.98, 0.95	0.99, 0.98	0.99, 0.998	0.01, 0.02
$f^5 7s^2 7p$	0.999, 0.89	0.999, 0.97	1.00, 0.97	0.001, 0.03
$f^5 6d7s7p$	0.99, 0.81	0.99, 0.93	1.00, 0.93	0.01, 0.07
$f^6 7s8s$	0.92, 1.00	0.996, 1.00	1.00, 0.88	0.003, 0.00
$f^6 6d^2$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
$f^4 6d^2 7s^2$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
$f^5 6d^2 7p$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
all levels	0.99, 0.90	0.998, 0.99	0.999, 0.90	0.002, 0.01

than that of the lowest-lying $5f^6 6d^2$ level. These same 15 levels and an additional six are predicted as belonging to this configuration by the category-learning network. The category learning network also predicts one even-parity unknown as belonging to the $5f^4 6d^2 7s^2$ configuration. This level, however, lies lower in energy than the lowest $5f^4 6d^2 7s^2$ level.

IV. CONCLUSION

In this paper we have investigated the application of counterpropagation neural networks to the classification of unknown energy levels of Cm II and Pu I. We have found that uniflow networks may be trained at nearly

TABLE XIV. Values of performance measures averaged over test and training sets for Pu I odd-parity energy levels. The first number of each pair is the value for the uniflow network and the second is that for the category-learning network.

Category	Sensitivity	Specificity	PPV	FAR
		(a) Test sets		
$f^5 6d7s^2$	0.92, 0.85	0.94, 0.90	0.92, 0.91	0.06, 0.10
$f^5 6d^2 7s$	0.62, 0.62	0.86, 0.86	0.63, 0.59	0.14, 0.14
$f^6 7s7p$	0.59, 0.67	0.84, 0.87	0.63, 0.63	0.16, 0.13
$f^7 7s$	0.00, 0.00	0.997, 0.997	0.00, 0.00	0.003, 0.003
$f^6 6d7p$	0.57, 0.71	0.989, 0.99	0.67, 0.71	0.011, 0.01
$f^5 6d7s8s$	0.00, 0.00	0.997, 0.997	0.00, 0.00	0.003, 0.003
all levels	0.73, 0.73	0.95, 0.95	0.76, 0.73	0.05, 0.05
		(b) Training sets		
$f^5 6d7s^2$	0.99, 0.92	0.99, 0.95	1.00, 0.96	0.01, 0.05
$f^5 6d^2 7s$	0.97, 0.87	0.99, 0.95	1.00, 0.87	0.01, 0.05
$f^6 7s7p$	0.96, 0.89	0.99, 0.96	0.997, 0.87	0.01, 0.04
$f^7 7s$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
$f^6 6d7p$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
$f^5 6d7s8s$	1.00, 1.00	1.00, 1.00	1.00, 1.00	0.00, 0.00
all levels	0.97, 0.90	0.99, 0.98	0.999, 0.90	0.01, 0.02

100%. Although category-learning networks are trained at a somewhat lower percentage, both types of networks give correct predictions for about 75% of test-set energy levels. Usually the same test-set levels are misclassified by each type of network although sometimes the predicted configurations are different. The mistakes made by uniflow networks in the learning phase of network operation are always exhibited as ambiguous outputs, indicating that one or more competitive layer PE's are taking responsibility for two or more energy levels belonging to different configurations. This is not a significant problem since the configurations of levels that give rise to ambiguous outputs are almost always correctly learned by the category-learning network. There is a slight tendency for the networks to make nonsensical configuration predictions; i.e., a level is sometimes assigned to a configuration whose lowest energy level is higher than the level in question. Category-learning networks exhibit this behavior more often than uniflow networks.

The most widely applicable remark that can be made concerning misclassified energy levels is that they almost always have isotope shifts which fall in the experimentally observed range of isotope shifts of the predicted configuration; i.e., they have isotope shifts which are consistent with the predicted configuration. This is probably not too surprising since the isotope shift is highly dependent on configuration. If, as in Pu I, there is a relatively large amount of configuration mixing, the isotope shifts for different configurations will overlap and the utility of the isotope shift in making configuration assignments will decrease. In slightly different terms, when configuration mixing occurs the network has greater difficulty learning the more complex or ambiguous "rules" for assigning configurations to energy levels. Despite these shortcomings in the performance of counterpropagation networks, we feel that they have performed admirably, particularly in regards to the Pu I system which exhibits an extremely complex electronic structure. Hopefully the classifications of the unknowns in this paper can serve as a useful starting point for further study using more conventional quantum-mechanical calculations.

Our goal in the present work has been to obtain a better qualitative understanding and characterization of network performance; for example, we would like to begin to be able to identify those cases or levels where a network is likely to fail and those where it is likely to succeed. Ultimately (in later work), we would like to be able to relate input-output weights of the various PE's comprising the network to the performance of the network. Hopefully the weights of the input and normalizing layer PE's will give us insight into the relative importance of each of the four features in determining the classification of a given level. Such information would be somewhat analogous to that obtained from principal component and other statistical analyses of other atomic systems [2–4]. In these analyses it is possible to calculate the relative importance of each feature in determining classifications of levels. The "weights" so obtained are averages over an entire data set; i.e., each feature has only one weight and this weight is a measure of the importance of the feature in determining the configurations of *all* energy levels in a data set. It should be possible to obtain feature weights (from input and normalizing layer PE weights) from a neural network which describe the importance of a feature in determining the configuration of a single given level. A neural network potentially contains more detailed information about the atomic system under study. In a similar manner, weights of competitive layer PE's should yield information about which PE is "firing" for a given input to the network. This should be very useful in determining why a network fails to achieve 100% training and in rationalizing network misclassifications quantitatively. Such arguments would carry the qualitative discussions of this paper one step further by casting them in a more precise and quantitative manner.

Finally, we would like to mention one further possible extension of this work. In principle it is possible in a uniflow network for more than one output layer PE to "fire" at the same time but with different strengths. When applied to the classification of atomic energy levels this would be an indication of configuration mixing. If configuration percentages were known for the training-set levels, a uniflow network could learn these percentages and use them to predict the extent of configuration mixing in unknowns.

[1] K. L. Peterson, Phys. Rev. A **41**, 2457 (1990).
[2] K. L. Peterson, D. L. Anderson, and M. L. Parsons, Phys. Rev. A **17**, 270 (1978).
[3] R. V. Lewis and K. L. Peterson, Phys. Rev. A **35**, 1119 (1987).
[4] R. V. Lewis and K. L. Peterson, Phys. Rev. A **38**, 3773 (1988).
[5] K. L. Peterson (unpublished).
[6] N. R. Bernstein and K. L. Peterson (unpublished).
[7] E. F. Worden, J. G. Conway, and J. Blaise, J. Opt. Soc. Am. B **3**, 1092 (1986).
[8] Available from NeuralWare, Inc., Penn Center West, Building IV, Suite 227, Pittsburgh, PA 15276.
[9] P. R. Gorman and T. J. Sejnowski, Neural Networks **1**, 75 (1988).
[10] G. T. Toussaint, IEEE Trans. Inf. Theory **IT-20**, 472 (1974).
[11] *Neural Network PC Tools, A Practical Guide,* edited by R. C. Eberhart and R. W. Dobbins (Academic, San Diego, 1990), pp. 169–174.
[12] J. Blaise, M. Fred, and R. G. Gutmacher, J. Opt. Soc. Am. B **3**, 403 (1986).