

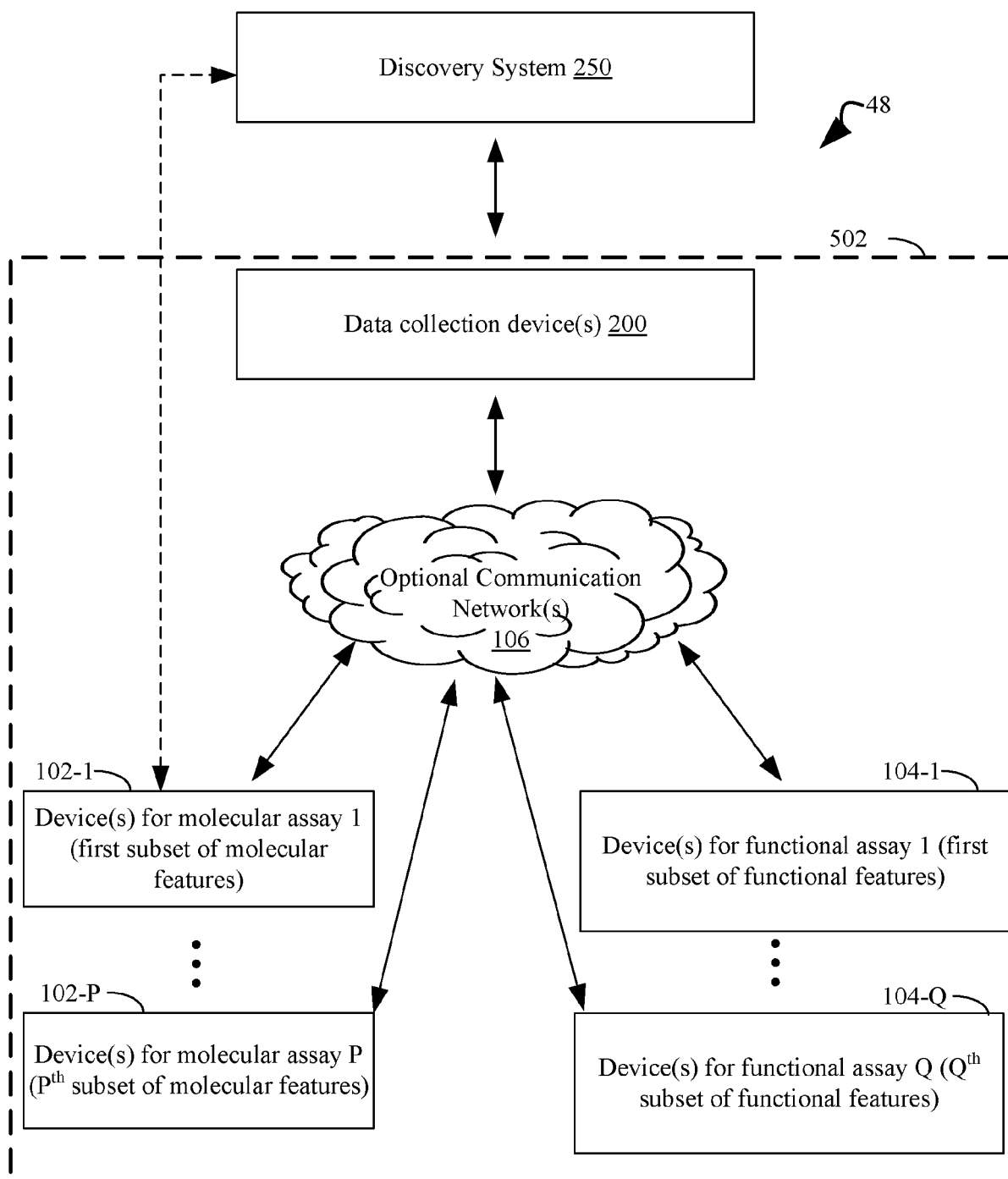


Fig. 1

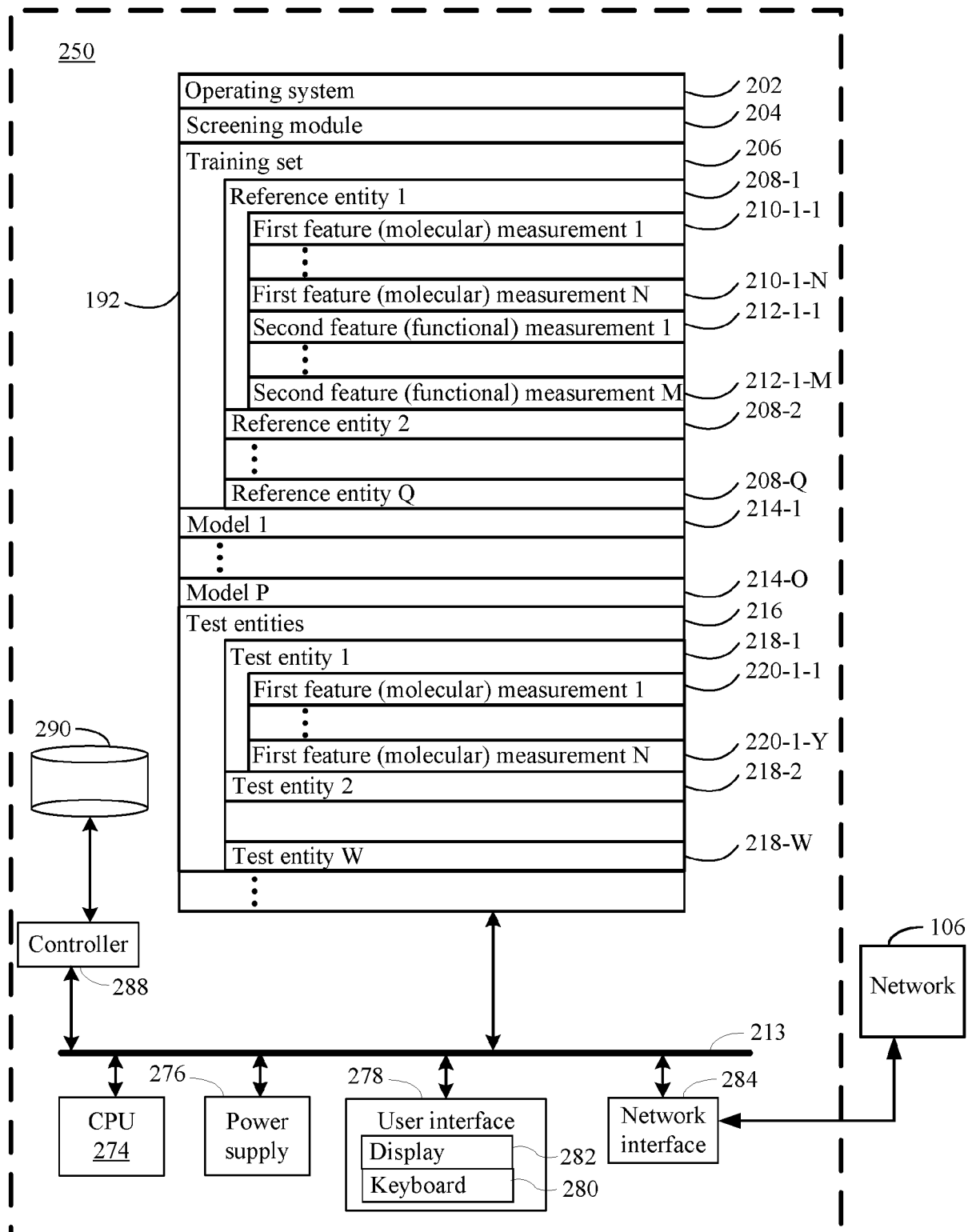


Fig. 2

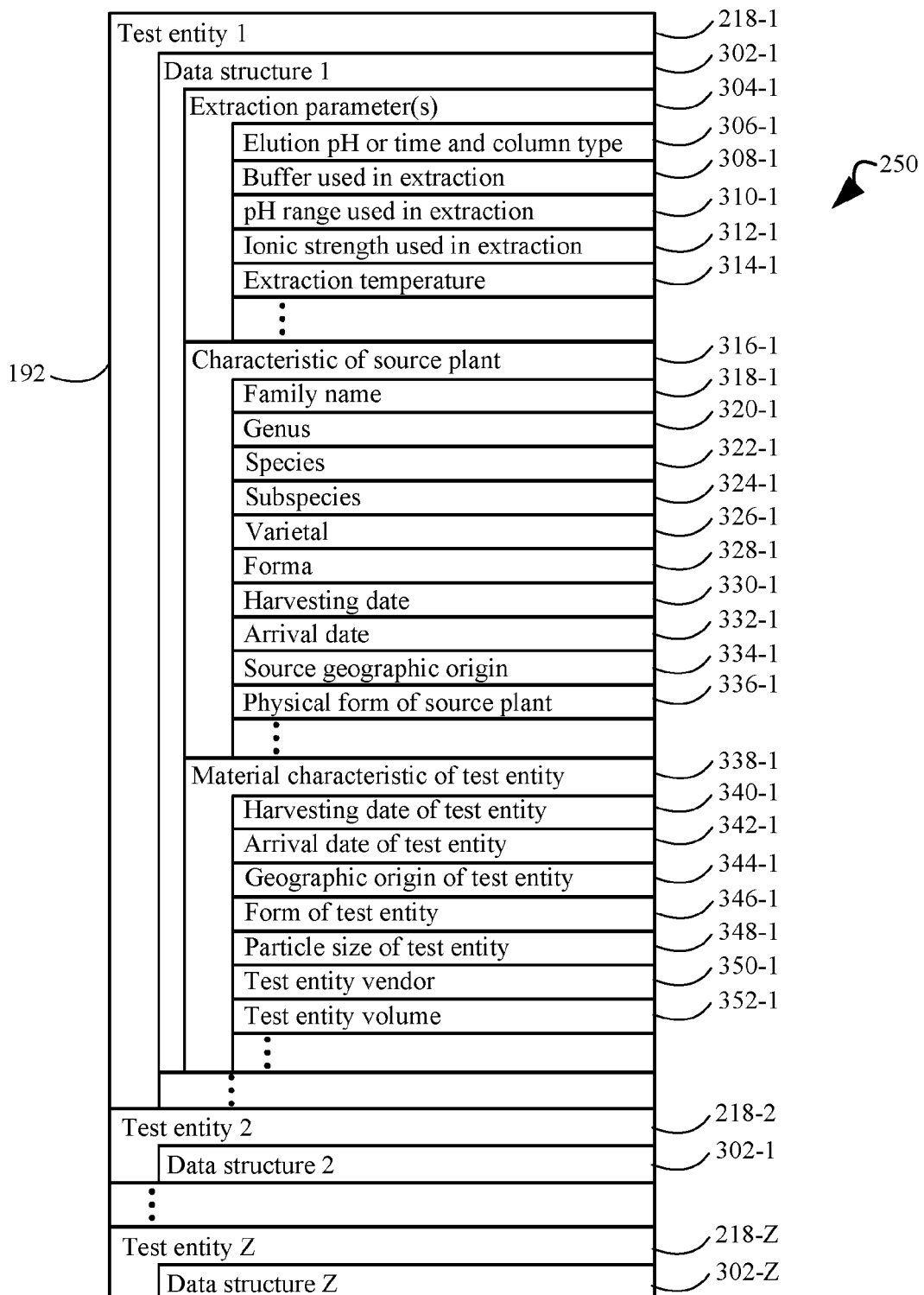


Fig. 3

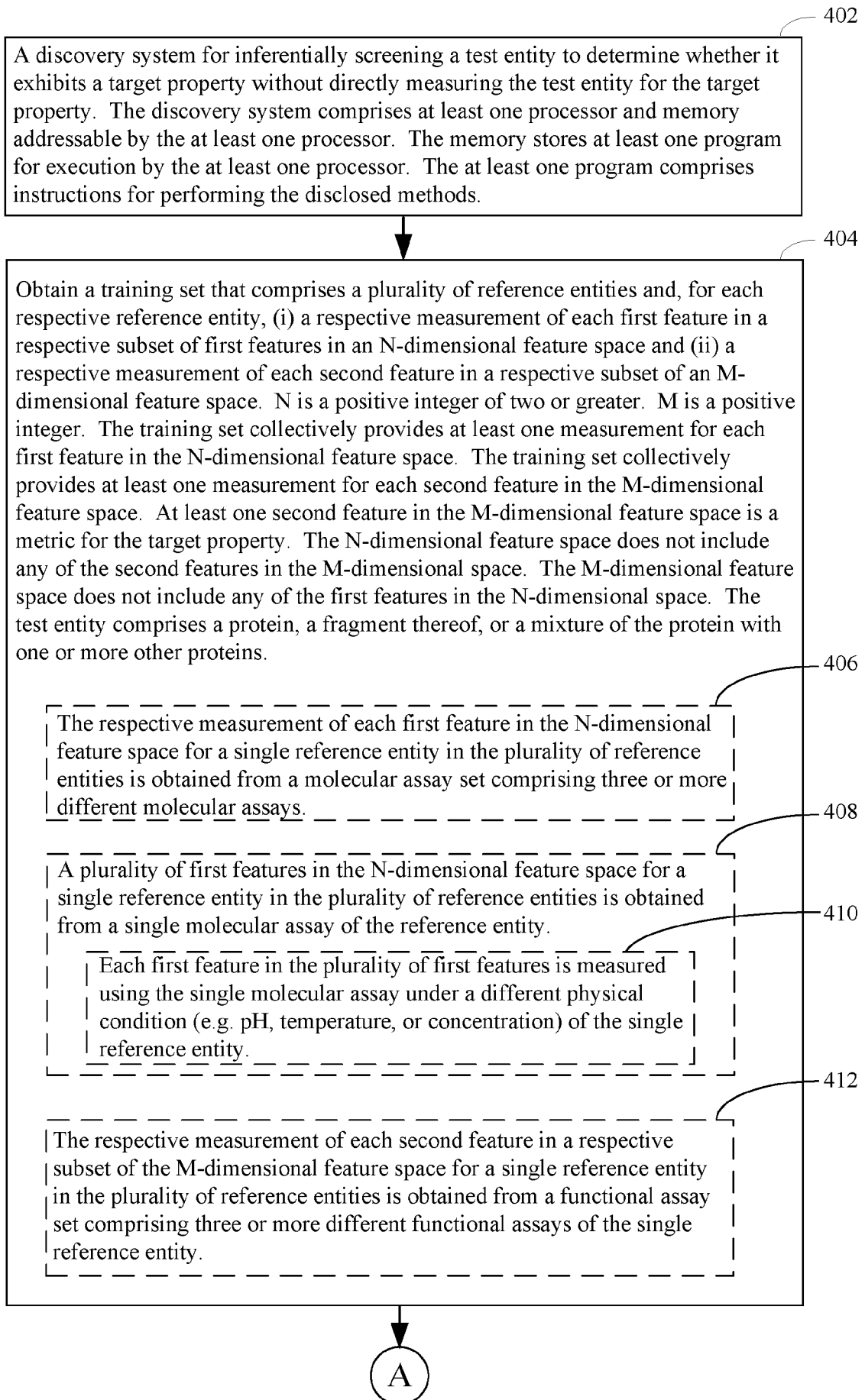


Fig. 4A

5/18

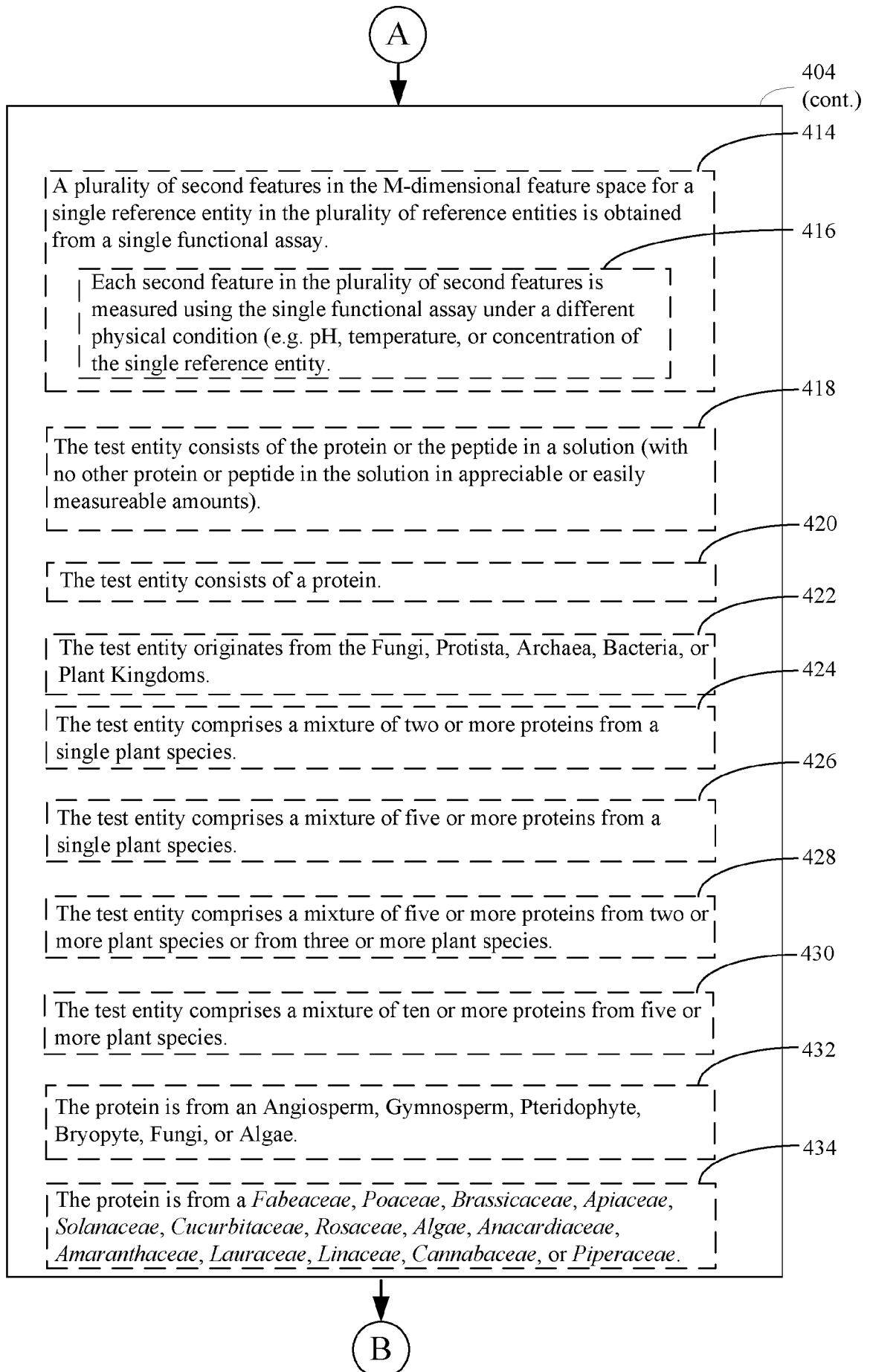


Fig. 4B

6/18

B

404
(cont.)

A first subset of the plurality of reference entities do not exhibit the target property or exhibit the target property to a lesser degree than a second subset of the plurality of reference entities.

A representation of the training set as a matrix of the plurality of reference entities against the N-dimensional feature space and the M-dimensional feature space has a sparseness of 80 percent or greater, 85 percent or greater, 90 percent or greater, 95 percent or greater, or 99 percent or greater.

Identify two or more first features, or one or more combinations thereof, in the N-dimensional feature space using a feature selection method and the training set, thereby selecting a set of first features $\{p_1, \dots, p_{N-K}\}$ from the N-dimensional feature space, where N-K is a positive integer less than N.

N is 10 or more and N-K is 5 or less

N is 35 or more, M is 60 or more, and N-K is 10 or less

N is 200 or more, and an absolute value of a coefficient of multiple correlation between measurements for the N-K first features across the training set and measurements for a single second feature across the training set is 0.60 or greater, 0.70 or greater, 0.80 or greater, or 0.90 or greater.

The feature selection method comprises regularization (e.g., is Lasso, LARS, or Elastic net) across the training set using the N-dimensional feature space and a single second feature in the M-dimensional feature space.

The feature selection method comprises application of a decision tree to the training set using the N-dimensional feature space and all or a portion of the M-dimensional feature space.

The feature selection method comprises application of Gaussian process regression to the training set using the N-dimensional feature space and a single second feature in the M-dimensional feature space.

C

Fig. 4C

7/18

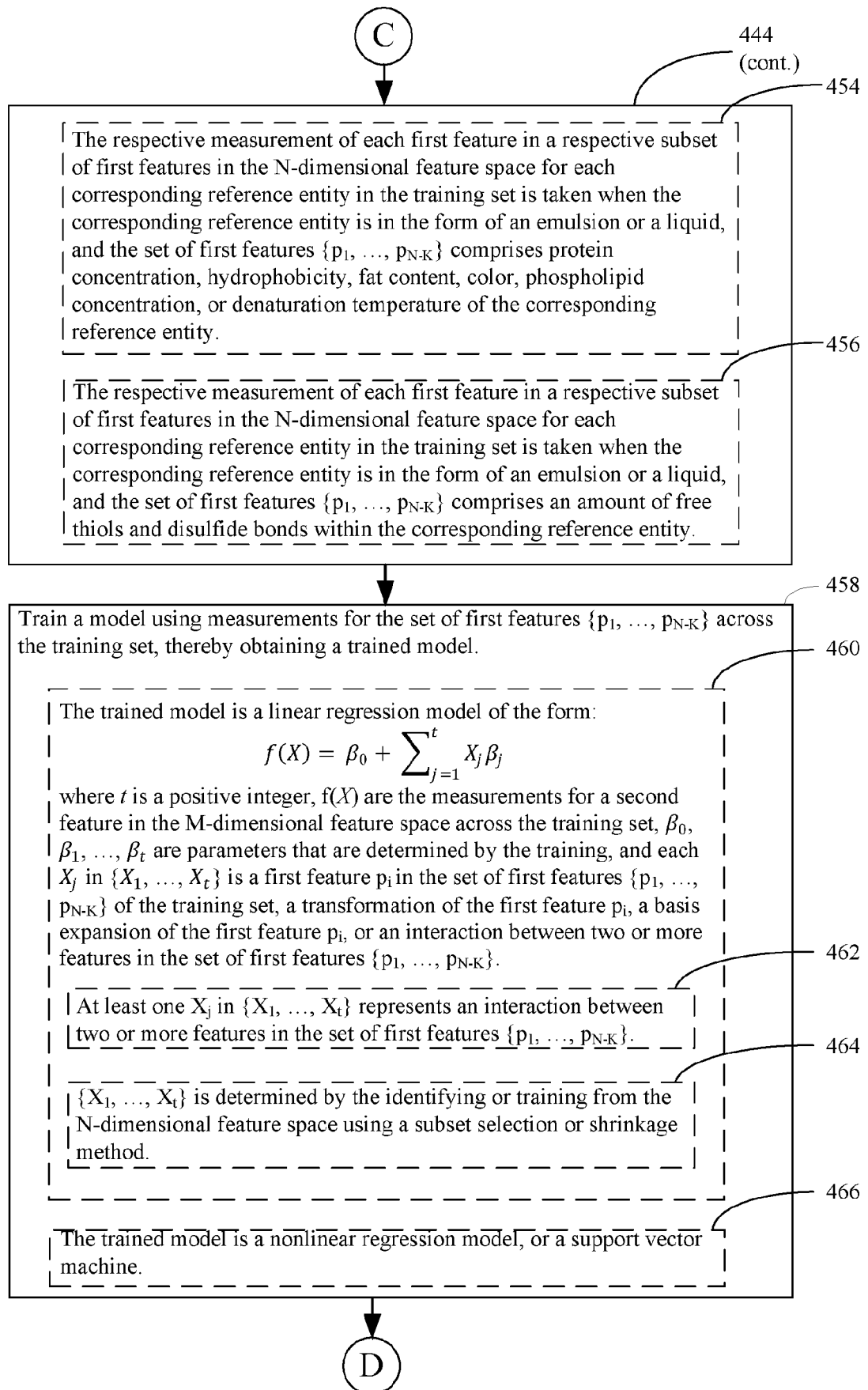


Fig. 4D

8/18



458 (cont.)

468

A first plurality of first features in the N-dimensional feature space is obtained from a first molecular assay of each reference entity in the training set. A second plurality of first features in the N-dimensional feature space is obtained from a second molecular assay of each reference entity in the training set. The feature selection method comprises: (i) application of a first kernel function to the respective measurement of each measured first feature in the first plurality of first features in the N-dimensional feature space for each reference entity in the plurality of reference entities, thereby deriving a first kernel matrix, (ii) application of a second kernel function to the respective measurement of each measured first feature in the second plurality of first features in the N-dimensional feature space for each reference entity in the plurality of reference entities, thereby deriving a second kernel matrix, and (iii) applying principal component analysis to the first kernel matrix and the second kernel matrix thereby identifying a plurality of principal components wherein the plurality of principal components collectively represent the set of first features $\{p_1, \dots, p_{N-K}\}$ from the N-dimensional feature space. The training the model using measurements for the set of first features $\{p_1, \dots, p_{N-K}\}$ across the training set comprises training the model using the plurality of principal components samples for each reference entity in the plurality of reference entities.

470

The model is a support vector machine.

472

A plurality of first features in the N-dimensional feature space is obtained from a molecular assay of each reference entity in the training set. The feature selection method comprises: (i) application of a kernel function to the respective measurement of each measured first feature in the plurality of first features in the N-dimensional feature space for each reference entity in the plurality of reference entities thereby deriving a kernel matrix, and (ii) applying principal component analysis to the kernel matrix thereby identifying a plurality of principal components wherein the plurality of principal components collectively represent the set of first features $\{p_1, \dots, p_{N-K}\}$ from the N-dimensional feature space. The training of the model (e.g., support vector machine) using measurements for the set of first features $\{p_1, \dots, p_{N-K}\}$ across the training set comprises training the model using the plurality of principal components samples for each reference entity in the plurality of reference entities.



Fig. 4E

9/18

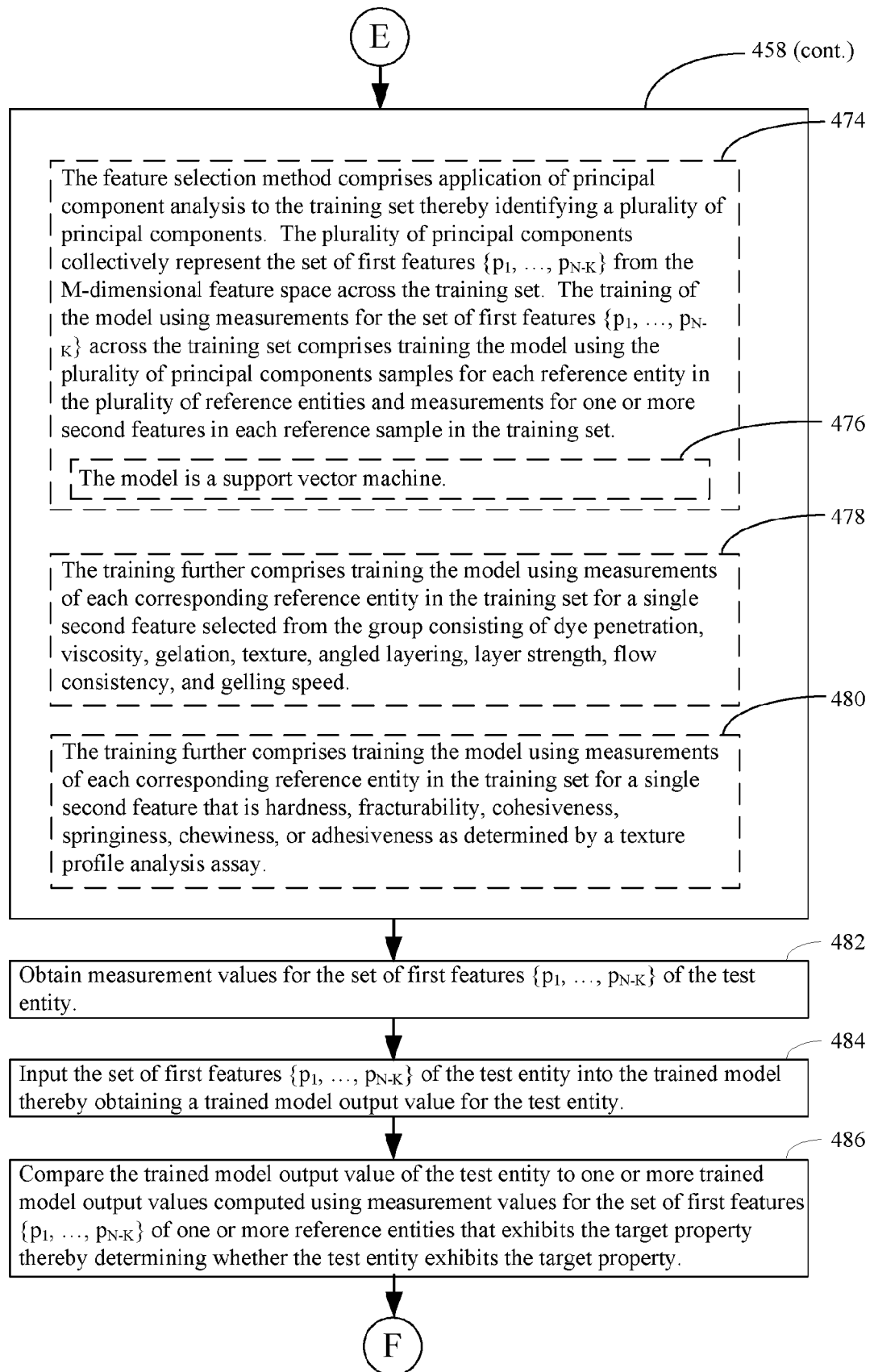


Fig. 4F

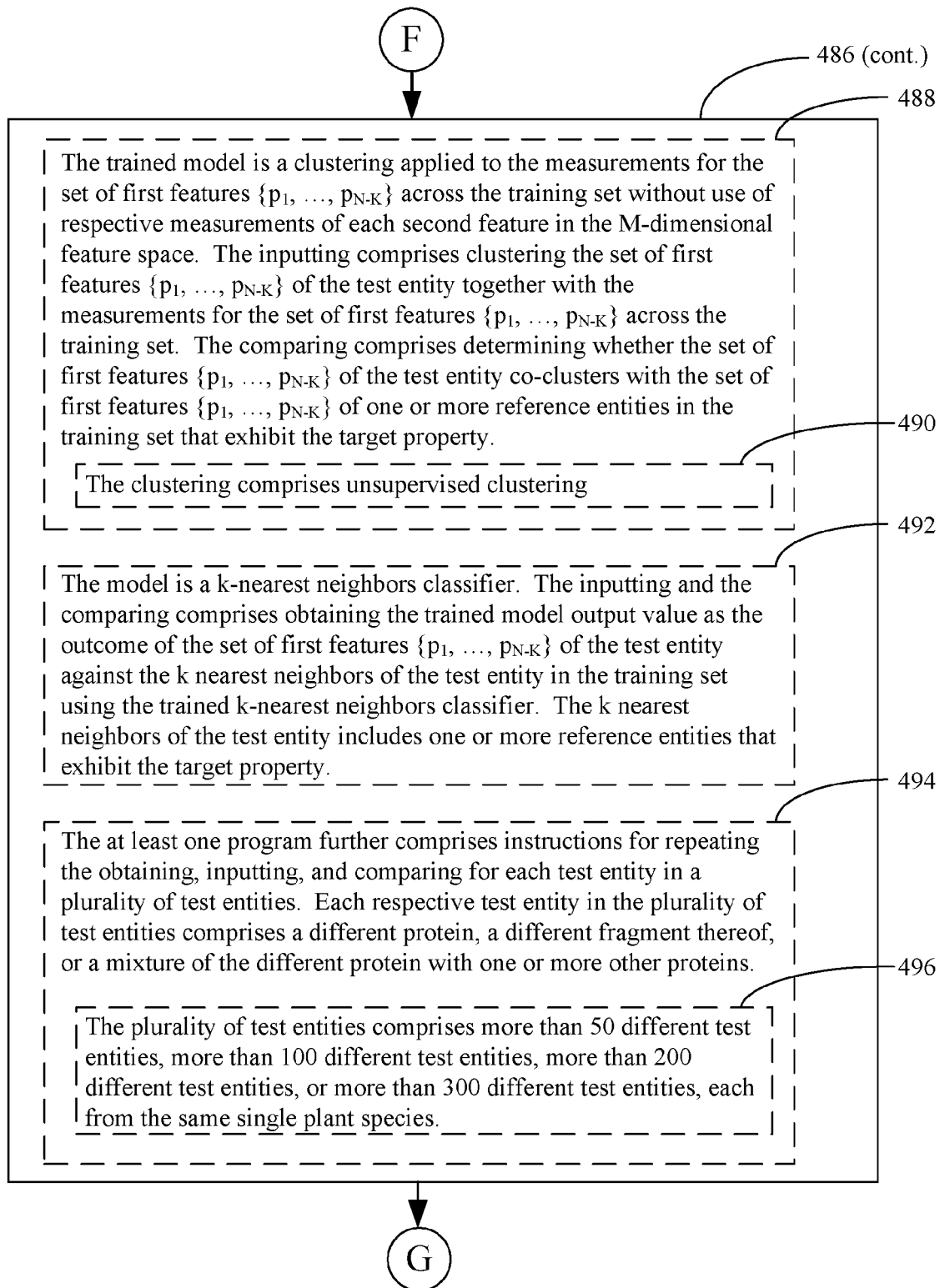


Fig. 4G

11/18

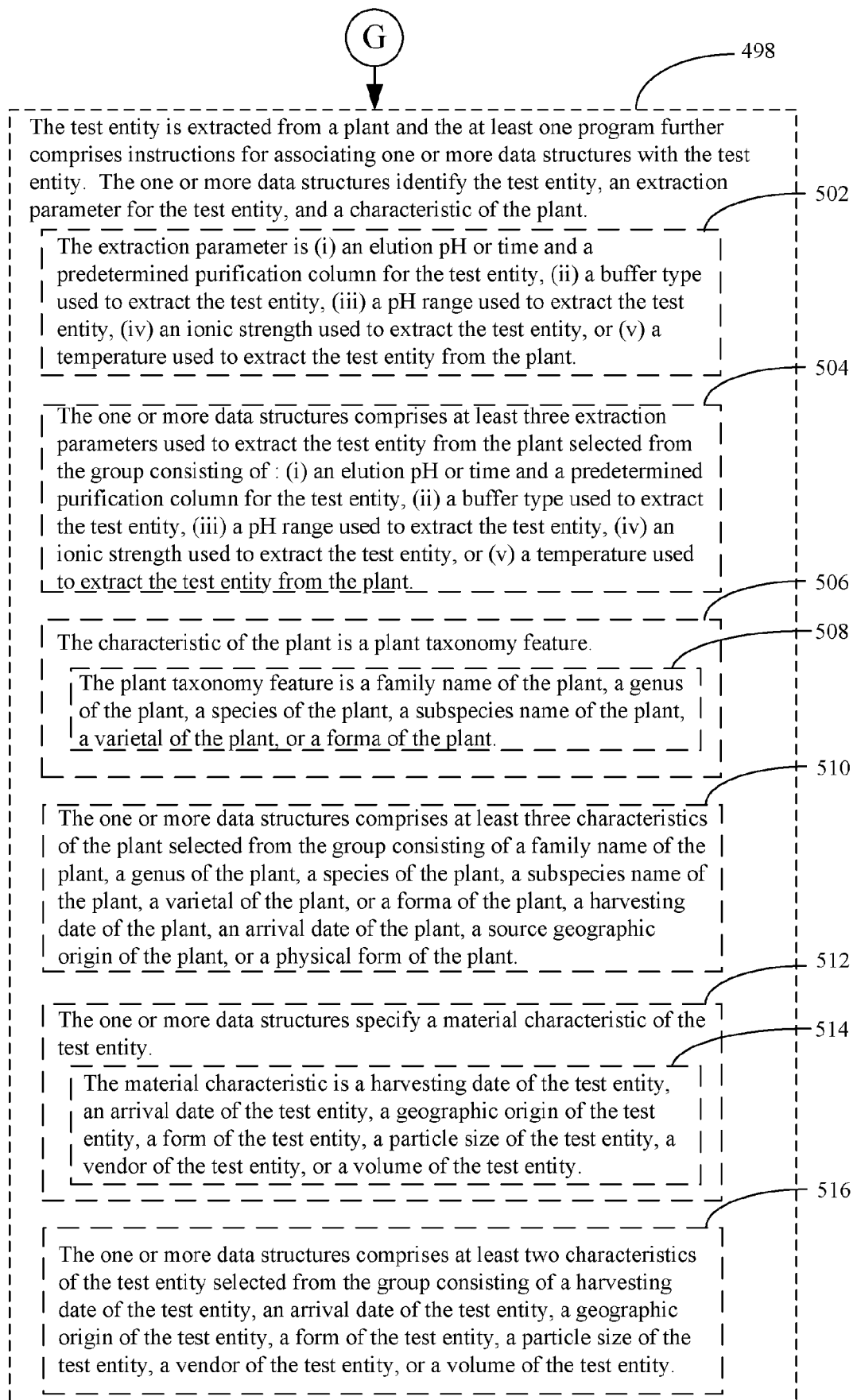


Fig. 4H

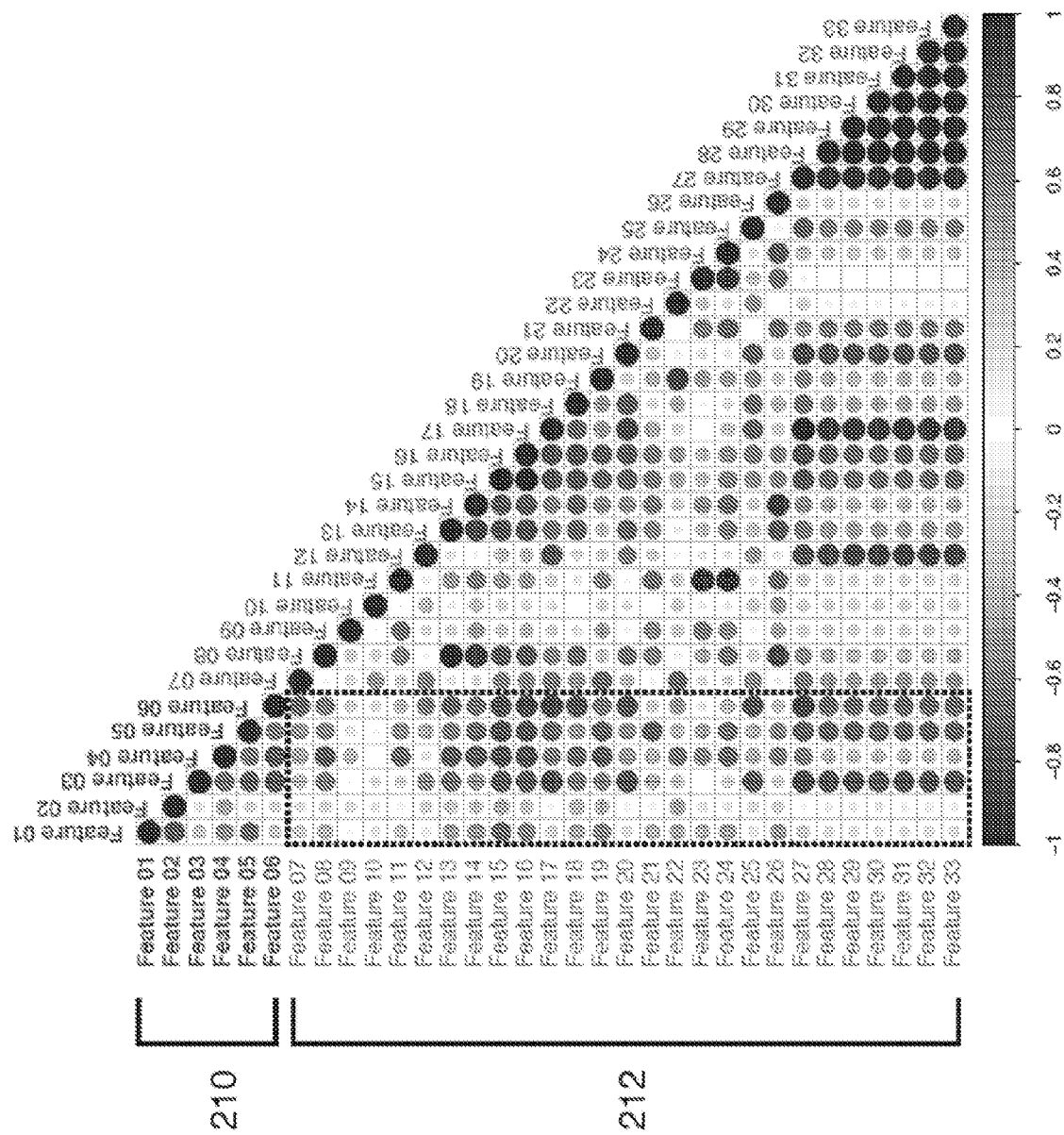


Fig. 5

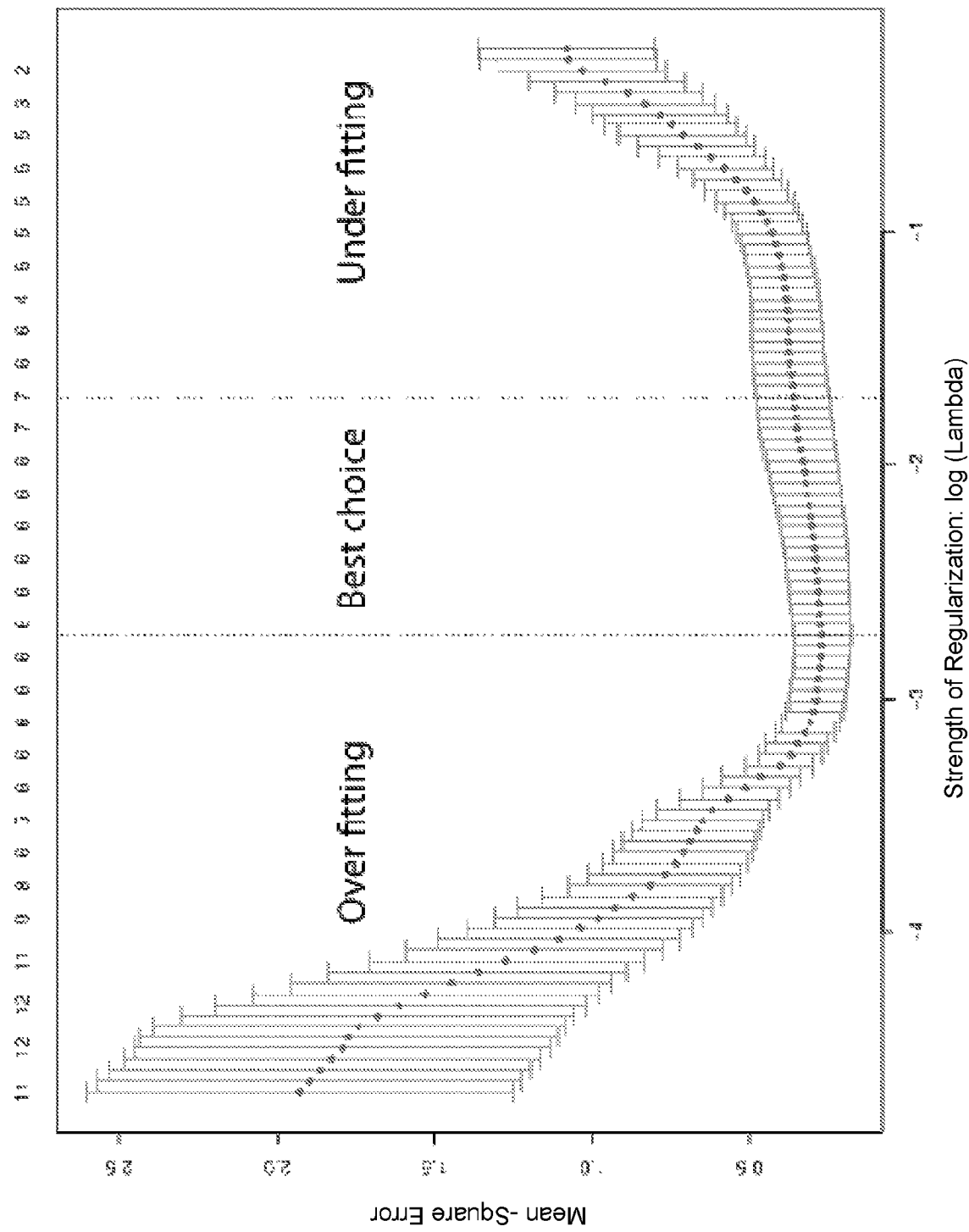


Fig. 6

Feature	Coefficients (λ @ min mse)	Coefficients (λ @ 1se)	p-value (of non-zero coefficients)
(Intercept)	0	0	-
Feature 1	1.54E-01	0	< 0.01
Feature 3	1.94E-01	1.86E-01	< 0.01
Feature 6	-3.22E-01	-2.67E-01	< 0.01
Feature 1:3:5	-4.25E-02	-2.06E-01	< 0.01
Feature 1:5:6	2.11E-01	1.13E-01	< 0.01
Feature 3:4:5	0	-3.50E-02	< 0.01
Feature 3:5:6	1.96E-01	7.76E-02	< 0.01

Fig. 7

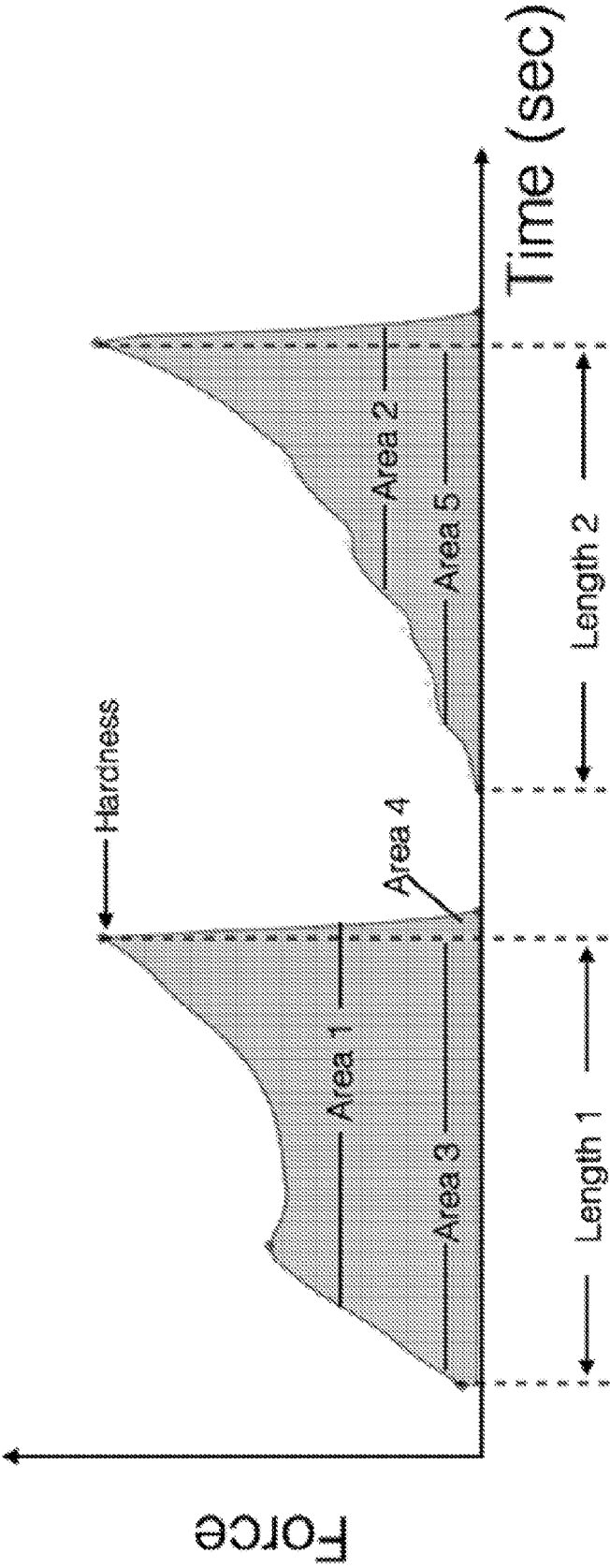


Fig. 8

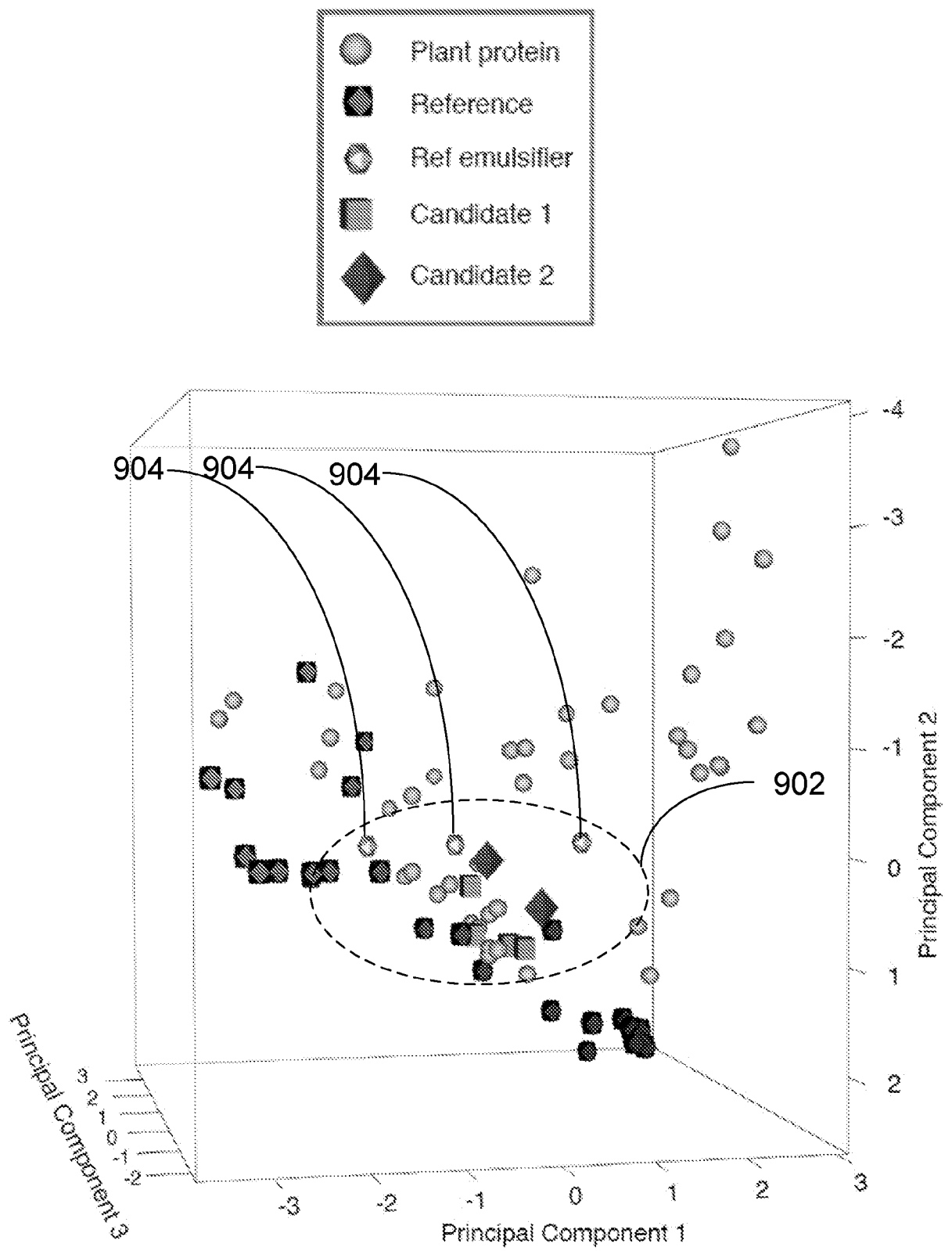


Fig. 9

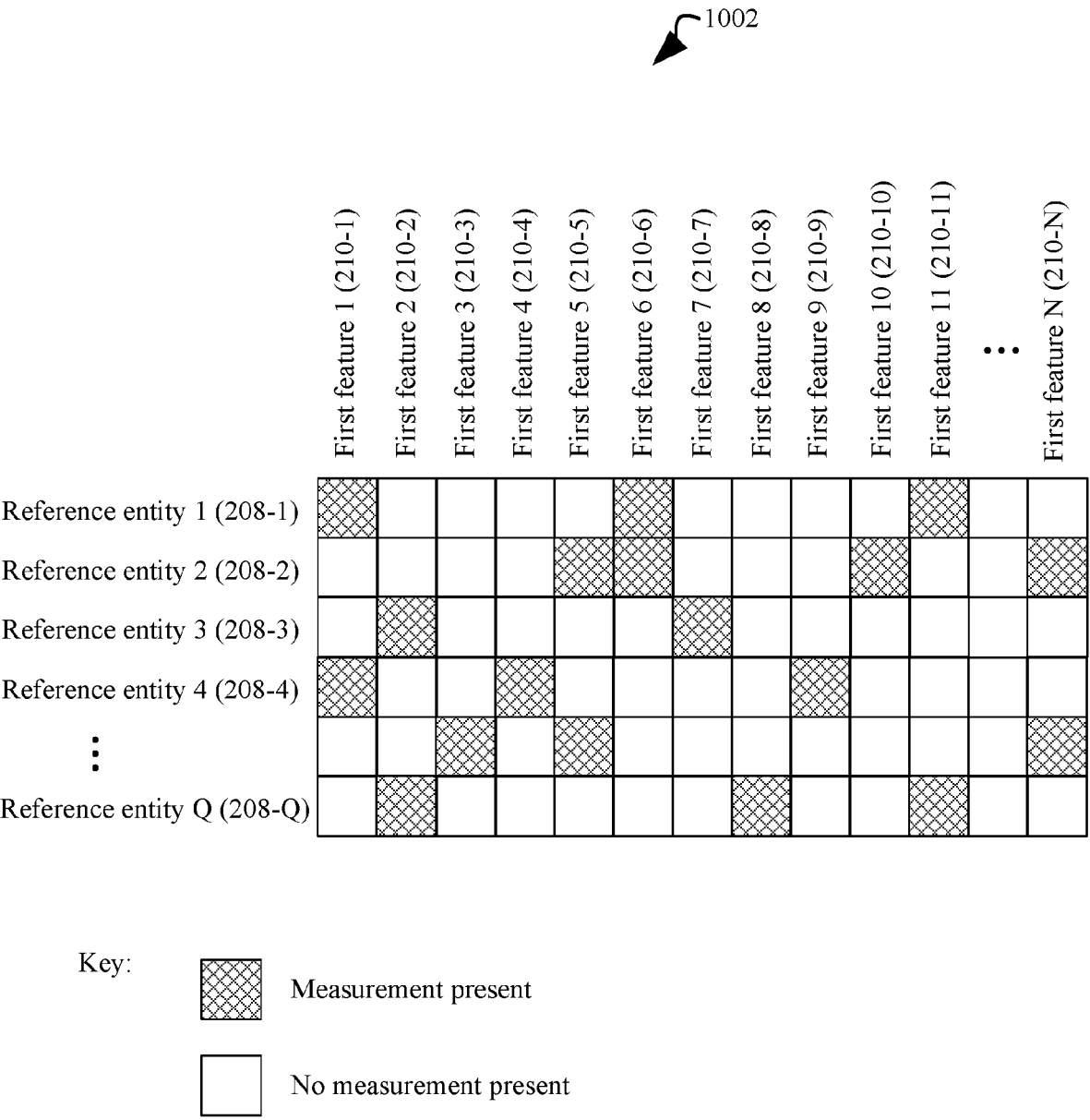


Fig. 10

1102 

	Second feature 1 (212-1)	Second feature 2 (212-2)	Second feature 3 (212-3)	Second feature 4 (212-4)	Second feature 5 (212-5)	Second feature 6 (212-6)	Second feature 7 (212-7)	Second feature 8 (212-8)	Second feature 9 (212-9)	Second feature 10 (212-10)	...	Second feature N (212-M)
Reference entity 1 (208-1)												
Reference entity 2 (208-2)												
Reference entity 3 (208-3)												
Reference entity 4 (208-4)												
Reference entity 5 (208-5)												
Reference entity 6 (208-6)												
Reference entity 7 (208-7)												
Reference entity 8 (208-8)												
Reference entity 9 (208-9)												
Reference entity 10 (208-10)												
Reference entity 11 (208-11)												
Reference entity 12 (208-12)												
Reference entity 13 (208-13)												
Reference entity 14 (208-14)												
Reference entity 15 (208-15)												
Reference entity 16 (208-16)												
...												
Reference entity Q (208-Q)												

Key:

 Measurement present

 No measurement present

Fig. 11