CLAIMS

1. A method of determining sequence information from two or more polynucleotide sequence portions, the method comprising:

(a) obtaining first intensity data comprising a combined intensity of a first signal obtained based upon a respective first nucleobase of at least one first polynucleotide sequence portion and a second signal obtained based upon a respective second nucleobase of at least one second polynucleotide sequence portion;

(b) obtaining second intensity data comprising a combined intensity of a third signal obtained based upon the respective first nucleobase of the at least one first polynucleotide sequence portion and a fourth signal obtained based upon the respective second nucleobase of the at least one second polynucleotide sequence portion;

(c) selecting one of a plurality of classifications based on the first and the second intensity data, wherein each classification of the plurality of classifications represents one or more possible combinations of respective first and second nucleobases, and wherein at least one classification of the plurality of classifications represents more than one possible combination of respective first and second nucleobases; and

(d) based on the selected classification, determining sequence information from the at least one first polynucleotide sequence portion and the at least one second polynucleotide sequence portion.

2. The method of claim 1, wherein the first and second signals and/or the third and fourth signals are obtained substantially simultaneously.

3. The method of claim 1 or claim 2, wherein selecting the classification based on the first and second intensity data comprises selecting the classification based on the combined intensity of the first and second signals and the combined intensity of the third and fourth signals.

4. The method of any one preceding claim, wherein, when based on a nucleobase of the same identity, an intensity of the first signal is substantially the same as an intensity of the second signal and an intensity of the third signal is substantially the same as an intensity of the fourth signal.

5. The method of any one preceding claim, wherein the plurality of classifications consists of a predetermined number of classifications.

6. The method of any one preceding claim, wherein the plurality of classifications comprises:

one or more classifications representing matching first and second nucleobases; and

one or more classifications representing mismatching first and second nucleobases, and

wherein determining sequence information from the at least one first polynucleotide sequence portion and the at least one second polynucleotide sequence portion comprises:

in response to selecting a classification representing matching first and second nucleobases, determining a match between the first and second nucleobases; or

in response to selecting a classification representing mismatching first and second nucleobases, determining a mismatch between the first and second nucleobases.

7. The method of claim 6, wherein determining sequence information from the at least one first polynucleotide sequence portion and the at least one second polynucleotide sequence portion comprises, in response to selecting a classification representing a match between the first and second nucleobases, base calling the first and second nucleobases.

8. The method of any one preceding claim, wherein determining sequence information from the at least one first polynucleotide sequence portion and the at least one second polynucleotide sequence portion comprises, based on the selected classification, determining that the second polynucleotide sequence portion is modified relative to the first polynucleotide sequence portion at a location associated with the first and second nucleobases.

9. The method of any one preceding claim, wherein at least one polynucleotide sequence comprises the first polynucleotide sequence portion and the second polynucleotide sequence portion.

10. The method of any one of claims 1 to 8, wherein at least one first polynucleotide sequence comprises the first polynucleotide sequence portion and at least one second polynucleotide sequence comprises the second polynucleotide sequence portion.

11. The method of any one preceding claim, wherein the first signal, second signal, third signal and fourth signal are generated based on light emissions associated with the respective nucleobase and detected at a sensor, and wherein the obtained signals are generated by:

contacting a plurality of polynucleotide molecules comprising the first and second polynucleotide sequence portions with first primers for sequencing the first polynucleotide sequence portion and second primers for sequencing the second polynucleotide sequence portion;

extending the first primers and the second primers by contacting the polynucleotide molecules with labeled nucleobases to form first labeled primers and second labeled primers;

stimulating the light emissions from the first and second labeled primers; and detecting the light emissions at a sensor.

12. The method of claim 11, wherein:

the first and second signals are based on light emissions detected in a first range of optical frequencies;

the third and fourth signals are based on light emissions detected in a second range of optical frequencies; and

wherein the first range of optical frequencies and the second range of optical frequencies are not identical.

13. The method of claim 11 or claim 12, wherein the polynucleotide molecules comprising the first and second polynucleotide sequence portions are attached to a substrate, optionally a flow cell, and wherein the light emissions from the first labeled primers and the light emissions from the second labeled primers are emitted from the same region or substantially overlapping regions of the substrate.

14. The method of any one of claims 11 to 13, wherein the light emissions detected at the sensor are spatially unresolved.

15. The method of claim 14, wherein the sensor is configured to provide a single output based upon the first and second signals.

16. The method of any one of claims 11 to 15, wherein the sensor comprises a single sensing element.

17. The method of any one preceding claim, wherein the at least one first polynucleotide sequence portion and the at least one second polynucleotide sequence portion are present in a cluster.

18. The method of any one preceding claim, wherein the one of the plurality of classifications is selected based on the first and the second intensity data using a Gaussian mixture model.

19. A data processing system comprising means for carrying out a method comprising:

(a) obtaining first intensity data comprising a combined intensity of a first signal obtained based upon a respective first nucleobase of at least one first polynucleotide sequence portion and a second signal obtained based upon a respective second nucleobase of at least one second polynucleotide sequence portion;

(b) obtaining second intensity data comprising a combined intensity of a third signal obtained based upon the respective first nucleobase of the at least one first polynucleotide sequence portion and a fourth signal obtained based upon the respective second nucleobase of the at least one second polynucleotide sequence portion;

(c) selecting one of a plurality of classifications based on the first and the second intensity data, wherein each classification of the plurality of classifications represents one or more possible combinations of respective first and second nucleobases, and wherein at least one classification of the plurality of classifications represents more than one possible combination of respective first and second nucleobases; and

(d) based on the selected classification, determining sequence information from the at least one first polynucleotide sequence portion and the at least one second polynucleotide sequence portion.

20. A computer-readable storage medium comprising instructions which, when executed by a processor, cause the processor to carry out a method comprising:

(a) obtaining first intensity data comprising a combined intensity of a first signal obtained based upon a respective first nucleobase of at least one first polynucleotide sequence portion and a second signal obtained based upon a respective second nucleobase of at least one second polynucleotide sequence portion;

(b) obtaining second intensity data comprising a combined intensity of a third signal obtained based upon the respective first nucleobase of the at least one first polynucleotide sequence portion and a fourth signal obtained based upon the respective second nucleobase of the at least one second polynucleotide sequence portion;

(c) selecting one of a plurality of classifications based on the first and the second intensity data, wherein each classification of the plurality of classifications represents one or more possible combinations of respective first and second nucleobases, and wherein at least one classification of the plurality of classifications represents more than one possible combination of respective first and second nucleobases; and

(d) based on the selected classification, determining sequence information from the at least one first polynucleotide sequence portion and the at least one second polynucleotide sequence portion.