CLAIMS

1. A neural processing core for a neural network, the neural processing core comprising:

a synaptic memory array comprising synaptic memory cells arranged in a plurality of memory cell rows and columns;

a plurality of first input activation lines that are each connected to a respective memory cell row, whereby each of the plurality of first input activation lines is configured to receive first input activation signals;

a plurality of control lines that are each connected to a respective memory cell column, whereby each of the plurality of control lines is configured to receive a column-wise control signal for controlling the synaptic memory cells of the respective memory cell column in a column-wise manner;

wherein each of the plurality of memory cell columns is associated with a first sensing line that is configured to output a first analog electrical signal, and

wherein each of the synaptic memory cells comprises:

a first access transistor having a gate connected to the control line associated with the memory cell column to which the synaptic memory cell belongs, the first access transistor being configured to be controlled by the column-wise control signal; and

a first memristor having a first end connected to a source node of the first access transistor and a second end connected to the first sensing line associated with the memory cell column, such that the first memristors of the synaptic memory cells in each memory cell column collectively form a memristor column.

2. The neural processing core according to claim 1, wherein each of the plurality of memory cell columns is associated with an additional sensing line that is configured to output an additional analog electrical signal, the neural processing core further comprising:

a first additional memristor having a first end connected to the source node of the first access transistor and a second end connected to the additional sensing line associated with the memory cell column, such that the first additional memristors of the synaptic memory cells in each memory cell column collectively form an additional memristor column.

3. The neural processing core according to claim 1, further comprising:

a plurality of second input activation lines that are each connected to a respective memory cell row, wherein each of the plurality of second input activation lines is configured to receive second input activation signals,

wherein each synaptic memory cell in a memory cell row is connected to the second input activation line associated with that memory cell row;

wherein each synaptic memory cell further comprises:

a second access transistor having a gate connected to the control line associated with the memory cell column to which the synaptic memory cell belongs, the second access transistor being configured to be controlled by the column-wise control signal received by the control line;

a second memristor having a first end connected to a source node of the second access transistor and a second end connected to the first sensing line associated with the memory cell column, such that the first and second memristors of the synaptic memory cells in each memory cell column collectively form the memristor column;

a second additional memristor having a first end connected to the source node of the second access transistor and a second end connected to the second end of the first additional memristor, such that the first and second additional memristors of the synaptic memory cells in each memory cell column collectively form the additional memristor column.

4. The neural processing core according to claim 1, wherein

for each of the plurality of memory cell rows, the first input activation line associated with the memory cell row is configured to receive the first input activation signal for the synaptic memory cells of the memory cell row in a row-wise manner,

for each of the plurality of memory cell columns, and for each of the synaptic memory cells of the synaptic memory array, the gate of the first access transistor of the synaptic memory cell connected to the control line is a control gate and the first access transistor further comprises a floating gate, wherein the source node of the first access transistor is connected to the first sensing line associated with the memory cell column which the synaptic memory cell belongs to and a drain node of the first access transistor is connected to the first input activation line associated with the memory cell row which the synaptic memory cell belongs to.

5. The neural processing core according to claim 2, wherein the neural processing core further comprises a peripheral sensing circuit connected to the first sensing lines and to the additional sensing lines, and the peripheral sensing circuit is configured to process the first analog electrical signals and the additional analog electrical signals from the plurality of memory cell columns, respectively, based on time multiplexing.

6. The neural processing core according to claim 5, wherein the peripheral sensing circuit configured to process the first analog electrical signals based on time multiplexing comprises processing the first analog electrical signals from the plurality of memory cell columns, respectively, in turn; and wherein the peripheral sensing circuit configured to process the additional analog electrical signals based on time multiplexing comprises processing the additional analog electrical signals from the plurality of memory cell columns, respectively, in turn.

7. The neural processing core according to claim 5, wherein the first analog electrical signals from the plurality of memory cell columns comprise first analog current signals, and the additional analog electrical signals from the plurality of memory cell columns comprise additional analog current signals, wherein the peripheral sensing circuit further comprises:

a first current-to-voltage converter configured to convert each of the first analog current signals from the plurality of memory cell columns, in turn, to a first analog voltage signal, and a second current-to-voltage converter configured to convert each of the additional analog current signals from the plurality of memory cell columns, in turn, to an additional analog voltage signal; and

an analog-to-digital converter (ADC) connected to the first and second current-to-voltage converters and configured to digitize the first analog voltage signal and the additional analog voltage signal received from the respective first and second current-to-voltage converters,

wherein the first and second current-to-voltage converters and the ADC are each shared amongst the plurality of memory cell columns for processing the plurality of first analog current signals and the additional analog current signals from the plurality of memory cell columns.

8. The neural processing core according to claim 5, wherein the first analog electrical signals from the plurality of memory cell columns comprise first analog current signals, and the

additional analog electrical signals from the plurality of memory cell columns comprise additional analog current signals, wherein the peripheral sensing circuit further comprises:

a plurality of current-to-voltage converter pairs, each current-to-voltage converter pair configured to convert each first and additional analog current from a corresponding group of memory cell columns of the plurality of memory cell columns, in turn, to a respective first and a respective additional analog voltage signal associated with the corresponding group of memory cell columns;

an analog multiplexer configured to select one output amongst outputs of the plurality of current-to-voltage converter pairs and forward the selected output; and

an ADC connected to the analog multiplexer and configured to digitize the first and the additional analog voltage signals from the selected output by the analog multiplexer,

wherein each of the plurality of current-to-voltage converter pairs is shared amongst the corresponding group of memory cell columns for processing the first and the additional analog current signals from the corresponding group of memory cell columns, and the ADC is shared amongst the plurality of memory cell columns.

9. The neural processing core according to claim 8, wherein the neural processing core is configured to control the current-to-voltage converter which produced the selected output to continue to operate past a column time-multiplexing cycle period based on the ADC taking longer than the column time-multiplexing cycle period to latch-in the first and the additional analog voltage signals from the selected output.

10. The neural processing core according to claim 5, wherein the first and the additional analog electrical signals from the plurality of memory cell columns respectively comprise first and additional analog voltage signals and the peripheral sensing circuit comprises:

an analog-to-digital converter (ADC) configured to digitize each of the first and additional analog voltage signals from the plurality of memory cell columns, in turn,

wherein the ADC is shared amongst the plurality of memory cell columns for processing the first and the additional analog voltage signals from the plurality of memory cell columns.

11. A method of operating a neural processing core for a neural network, the neural processing core comprising:

a synaptic memory array comprising synaptic memory cells arranged in a plurality of memory cell rows and columns;

a plurality of first input activation lines that are each connected to a respective memory cell row, whereby each of the plurality of first input activation lines is configured to receive first input activation signals;

a plurality of control lines that are each connected to a respective memory cell column, whereby each of the plurality of control lines is configured to receive a column-wise control signal;

wherein each of the plurality of memory cell columns is associated with a first sensing line that is configured to output a first analog electrical signal, and

wherein each of the synaptic memory cells comprises:

a first access transistor having a gate connected to the control line associated with the memory cell column to which the synaptic memory cell belongs, the first access transistor being configured to be controlled by the column-wise control signal;

a first memristor having a first end connected to a source node of the first access transistor and a second end connected to the first sensing line associated with the memory cell column, such that the first memristors of the synaptic memory cells in each memory cell column collectively form a memristor column;

for performing inference on the synaptic memory array, the method further comprises:

applying a predetermined ground voltage to the first sensing line connected to the selected memory cell column; and

sending first input activation signals to the plurality of first input activation lines associated with the plurality of memory cell rows, respectively, and

for performing a write operation on a selected synaptic memory cell of the synaptic memory array, the method further comprises:

sending a column-wise control signal to the control line associated with the memory cell column which the selected synaptic memory cell belongs to;

applying a predetermined ground voltage to the first sensing line associated with the memory cell column which the selected synaptic memory cell belongs to;

sending a first programming signal as the first input activation signal to the first input activation line associated with the memory cell row which the selected synaptic memory cell belongs to; and

applying a non-state changing voltage or floating each of the first sensing lines associated with memory cell columns to which the selected synaptic memory cell does not belong to.

12. The method according to claim 11, wherein each of the plurality of memory cell columns is associated with an additional sensing line that is configured to output an additional analog electrical signal, wherein each of the synaptic memory cells further comprises:

a first additional memristor having a first end connected to the source node of the first access transistor and a second end connected to the additional sensing line associated with the memory cell column, such that the first additional memristors of the synaptic memory cells in each memory cell column collectively form an additional memristor column;

wherein for the step of performing the inference on the synaptic memory array, the method further comprises:

applying a predetermined ground voltage to the additional sensing line connected to the selected memory cell column;

wherein for performing a write operation on a selected synaptic memory cell of the synaptic memory array, the method further comprises:

applying a predetermined ground voltage to the additional sensing line associated with the memory cell column which the selected synaptic memory cell belongs to; and

applying a non-state changing voltage or floating each of the additional sensing lines associated with memory cell columns to which the selected synaptic memory cell does not belong to.

13. The method according to claim 12, wherein the neural processing core further comprises:

a plurality of second input activation lines that are each connected to a respective memory cell row, wherein each of the plurality of second input activation lines is configured to receive second input activation signals,

wherein each synaptic memory cell in a memory cell row is connected to the second input activation line associated with that memory cell row;

wherein each synaptic memory cell further comprises:

a second access transistor having a gate connected to the control line associated with the memory cell column to which the synaptic memory cell belongs, the second

access transistor being configured to be controlled by the column-wise control signal received by the control line;

a second memristor having a first end connected to a source node of the second access transistor and a second end connected to the first sensing line associated with the memory cell column, such that the first and second memristors of the synaptic memory cells in each memory cell column collectively form the memristor column;

a second additional memristor having a first end connected to the source node of the second access transistor and a second end connected to the second end of the first additional memristor, such that the first and second additional memristors of the synaptic memory cells in each memory cell column collectively form the additional memristor column;

for said performing inference on the synaptic memory array, the method further comprises:

sending second input activation signals to the plurality of second input activation lines associated with the plurality of memory cell rows, respectively, and

for said performing a write operation on a selected synaptic memory cell of the synaptic memory array, the method further comprises:

sending a second programming signal as the second input activation signal to the second input activation line associated with the memory cell row which the selected synaptic memory cell belongs to,

wherein one of the first and second programming signals is a programming signal for setting a conductance state of the corresponding memristor of the first and second memristors of the selected synaptic memory cell and the other one of the first and second programming signals is the predetermined ground voltage or no signal with the corresponding one of the first and second input activation lines floated, based on a polarity of a synaptic weight value to be stored by the selected synaptic memory cell.

14. The method according to claim 13,

wherein for said performing inference on the selected synaptic memory array, the method further comprises:

applying a predetermined ground voltage to the first sensing line and the additional sensing line connected to the selected memory cell column to which the selected synaptic memory array belongs to, and

wherein for said performing a write operation on the selected synaptic memory cell of the synaptic memory array, the write operation is performed on a selected first memristor and a selected second memristor and the one or more additional pairs of additional memristors of the selected synaptic memory cell, the method comprises:

said applying a predetermined ground voltage is applying the predetermined ground voltage to the first sensing line and the additional sensing line associated with the memristor column to which the selected memristors belong to, and

the method further comprises:

applying a non-state changing voltage or floating each of the first sensing lines and the additional sensing lines associated with memory cell columns to which the selected synaptic memory cell does not belong to.

15. The method according to claim 12, wherein

the neural processing core further comprises a peripheral sensing circuit connected to the first sensing lines and to the additional sensing lines, and the peripheral sensing circuit is configured to process the first analog electrical signals and the additional analog electrical signals from the plurality of memory cell columns, respectively, based on time multiplexing, and

for said performing inference on the synaptic memory array, the method further comprises processing, using the peripheral sensing circuit, the first analog electrical signals from the plurality of memory cell columns, respectively, based on time multiplexing.

16. The method according to claim 15, wherein

the first analog electrical signals from the plurality of memory cell columns comprise first analog current signals, and the additional analog electrical signals from the plurality of memory cell columns comprise additional analog current signals, wherein the peripheral sensing circuit further comprises:

a first current-to-voltage converter configured to convert each of the first analog current signals from the plurality of memory cell columns, in turn, to a first analog voltage signal, and a second current-to-voltage converter configured to convert each of the additional analog current signals from the plurality of memory cell columns, in turn, to an additional analog voltage signal; and

an analog-to-digital converter (ADC) connected to the first and second current-to-voltage converters and configured to digitize the first analog voltage signal and the additional analog voltage signal received from the respective first and second current-to-voltage converters,

wherein the first and second current-to-voltage converters and the ADC are each shared amongst the plurality of memory cell columns for processing the plurality of first analog current signals and the additional analog current signals from the plurality of memory cell columns, and

for said performing inference on the synaptic memory array, the method further comprises:

converting, using the first current-to-voltage converter, each of the plurality of first analog current signals from the plurality of memory cell columns, in turn, to a first analog voltage signal;

converting, using the second current-to-voltage converter, each of the plurality of additional analog current signals from the plurality of memory cell columns, in turn, to a second analog voltage signal;

digitizing, using the ADC, the first analog voltage signal received from the first current-to-voltage converter, and the additional analog voltage signal received from the second current-to-voltage converter.