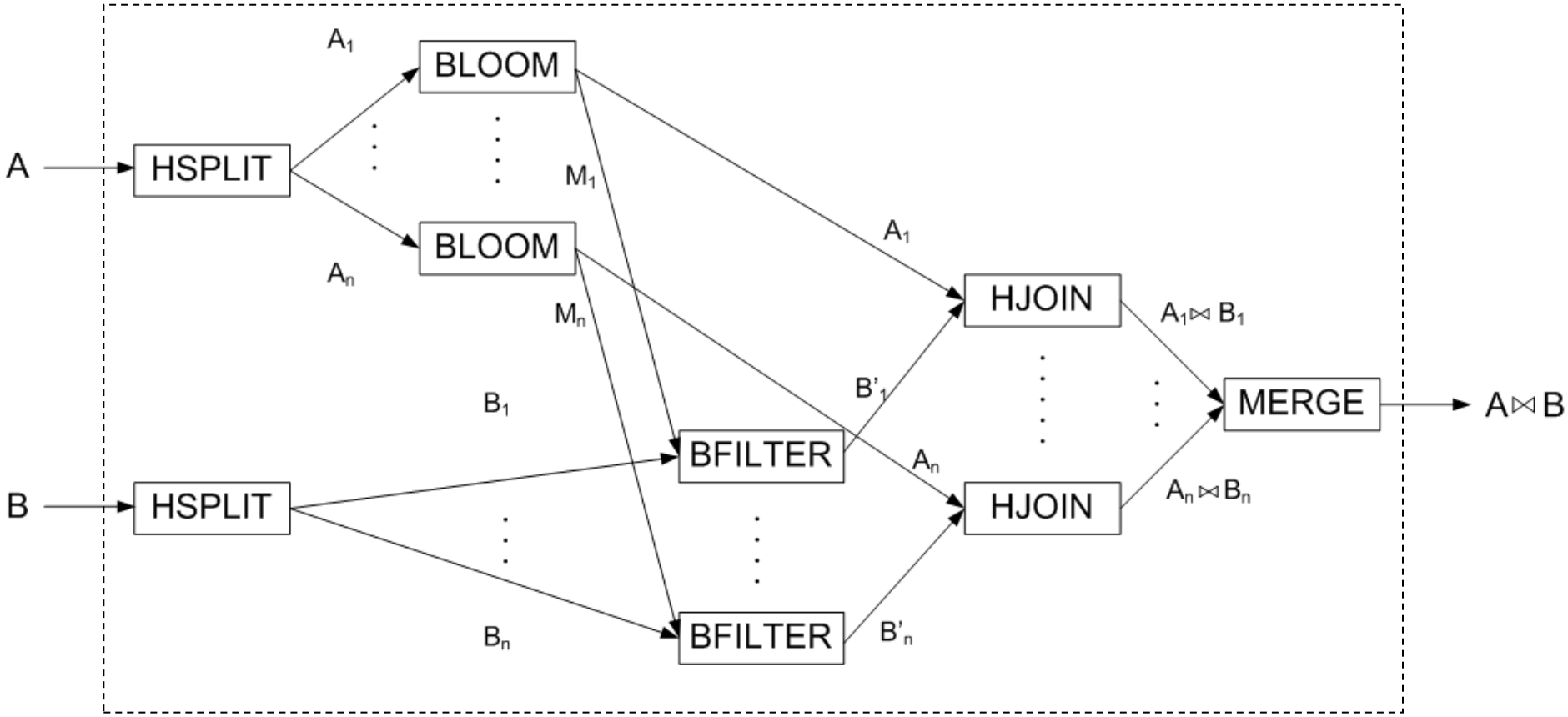
Illustration of join in Gamma. Two streams of tuples A and B are joined.

To be more efficient, the HJOIN can be implemented as follows (next page)



HSPLIT

- splits tuples of streams A and B to substreams A_1 to A_n and B_1 to B_n, respectively
- passes substreams of A to BLOOM boxes

BLOOM:

- clear bit map M
- read each A tuple, hash its join key, and mark corresponding bit in M

- output each tuple A
- after all A tuples read, output M
- there a n BLOOM boxes, i.e., as many as there are substreams of A

BFILTER

- read bit map M
- read each tuple of B, hash its join key: if corresponding bit in M is not set discard tuple (as it will never join with A tuples)
- else output tuple in substreams from B'_1 to B'_n
- there a n BFILTER boxes, i.e., as many as there are substreams of B

HJOIN

- joins tuples of A and B'
- there a n HJOIN boxes, i.e., as many as there are substreams of A and B

MERGE

- merges the result of all HJOIN boxes
- outputs the joined stream of tuples A and B