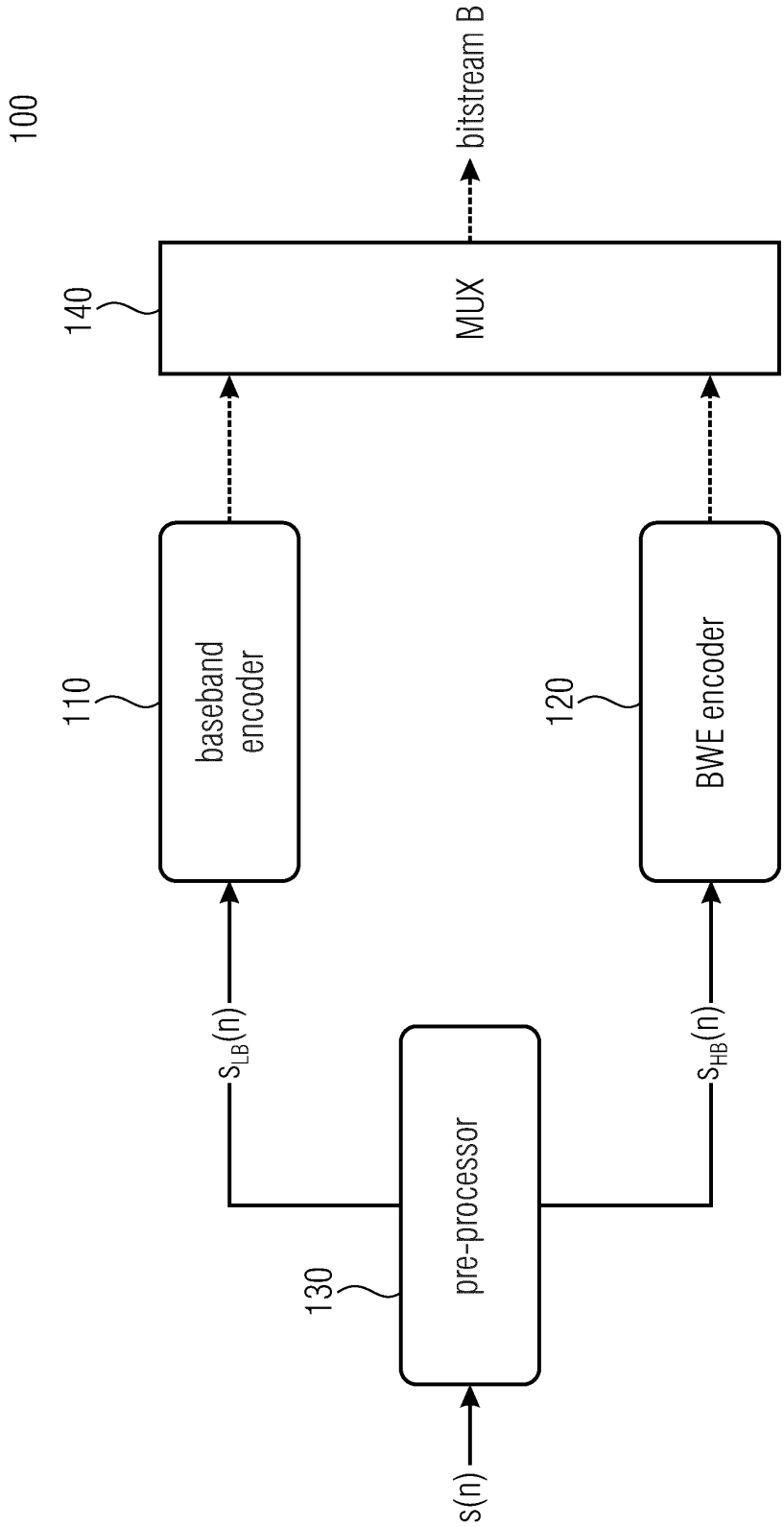

100
bitstream B
MUX
140
baseband encoder
110
BWE encoder
120
$s_{LB}(n)$
$s_{HB}(n)$
pre-processor
130
s(n)

Fig. 1

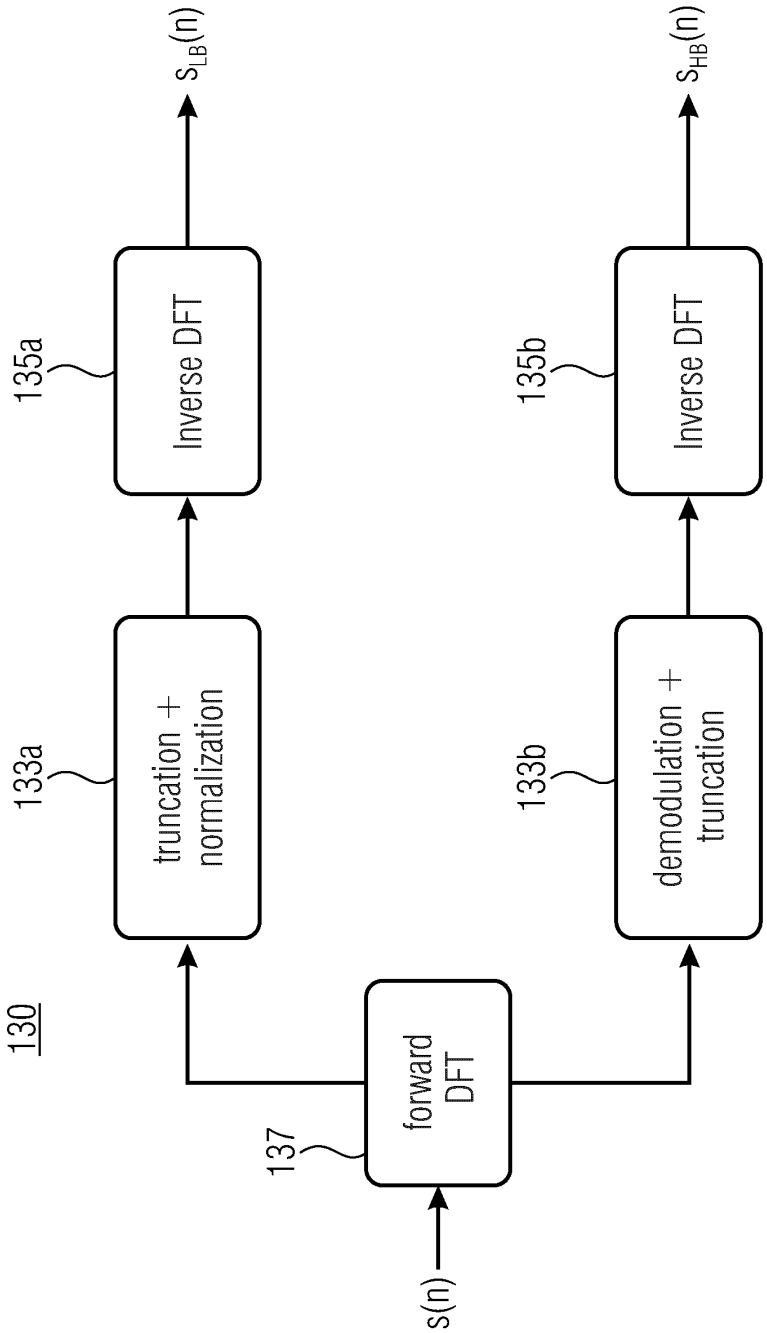

130
s(n)
137
forward DFT
133a
truncation + normalization
135a
Inverse DFT
$s_{LB}(n)$
133b
demodulation + truncation
135b
Inverse DFT
$s_{HB}(n)$

Fig. 2

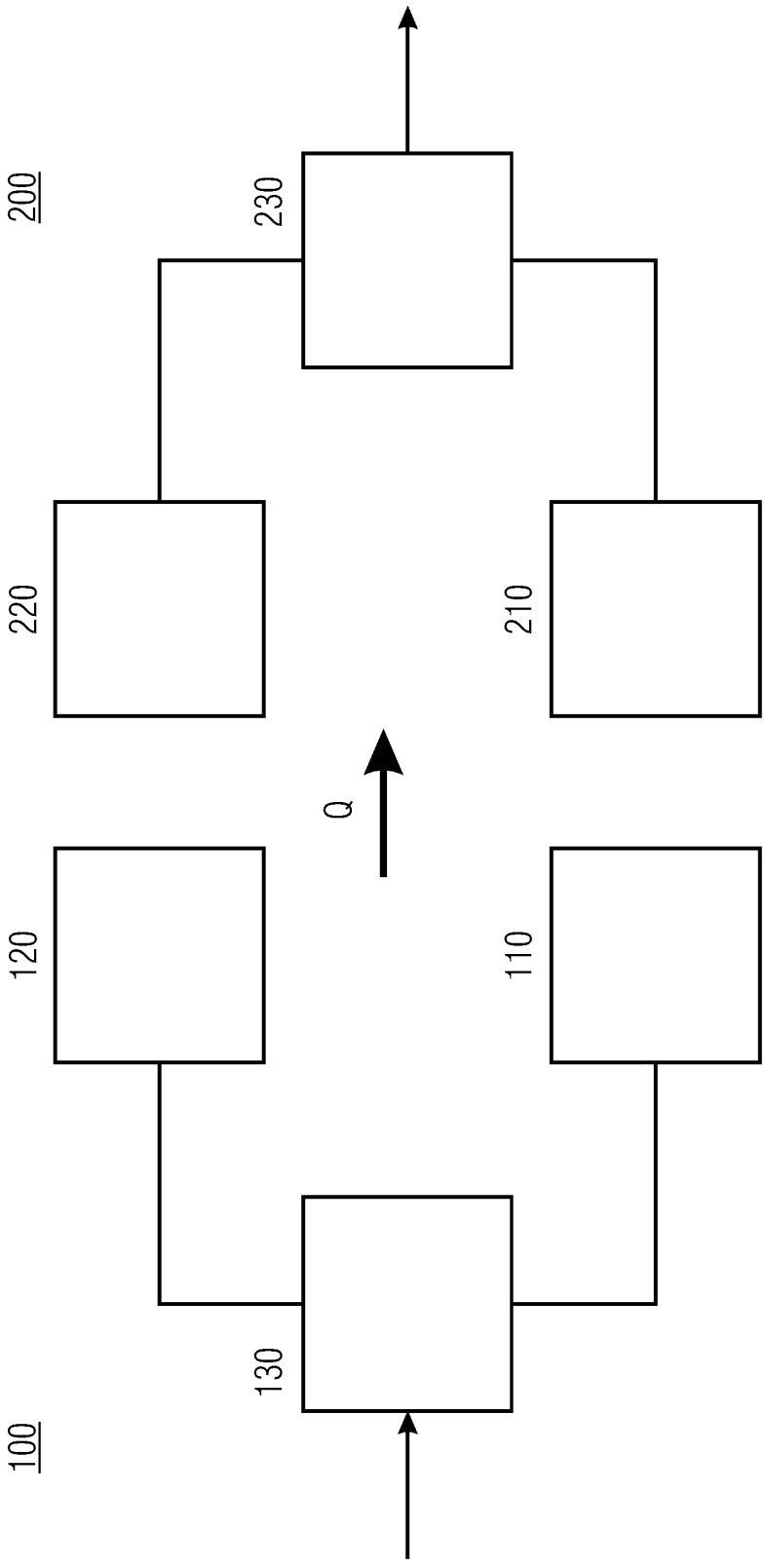


Fig. 3

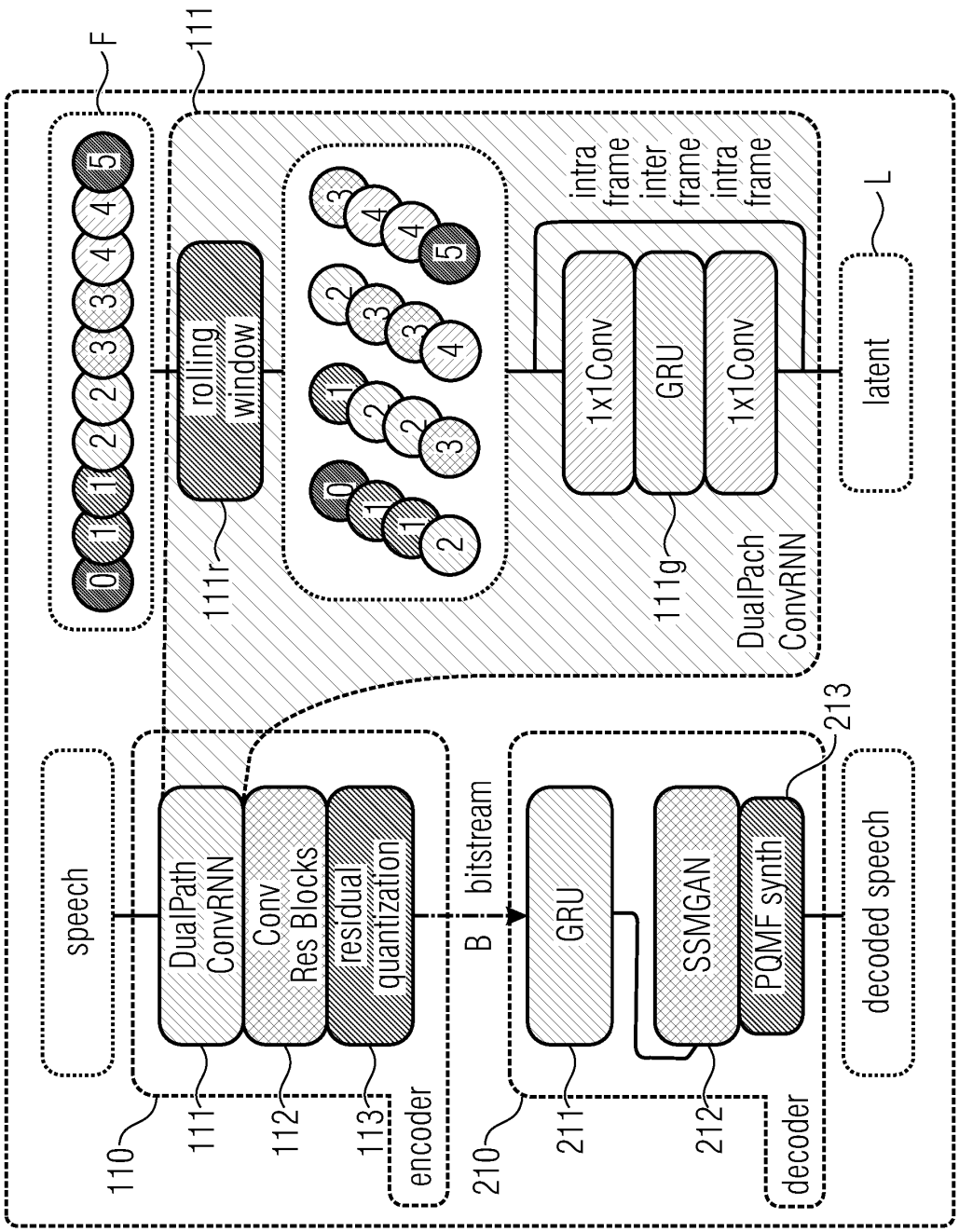

F
111
0 1 1 2 2 3 3 4 4 5
rolling window
111r
0 1 1 2 2 3 1 2 3 3 4 2 3 4 4 5
intra frame
inter frame
intra frame
1x1Conv
GRU
1x1Conv
111g
DualPach ConvRNN
L
latent
speech
DualPath ConvRNN
Conv Res Blocks
residual quantization
B bitstream
GRU
SSMGAN
PQMF synth
213
decoded speech
110
111
112
113
encoder
210
211
212
decoder

Fig. 4

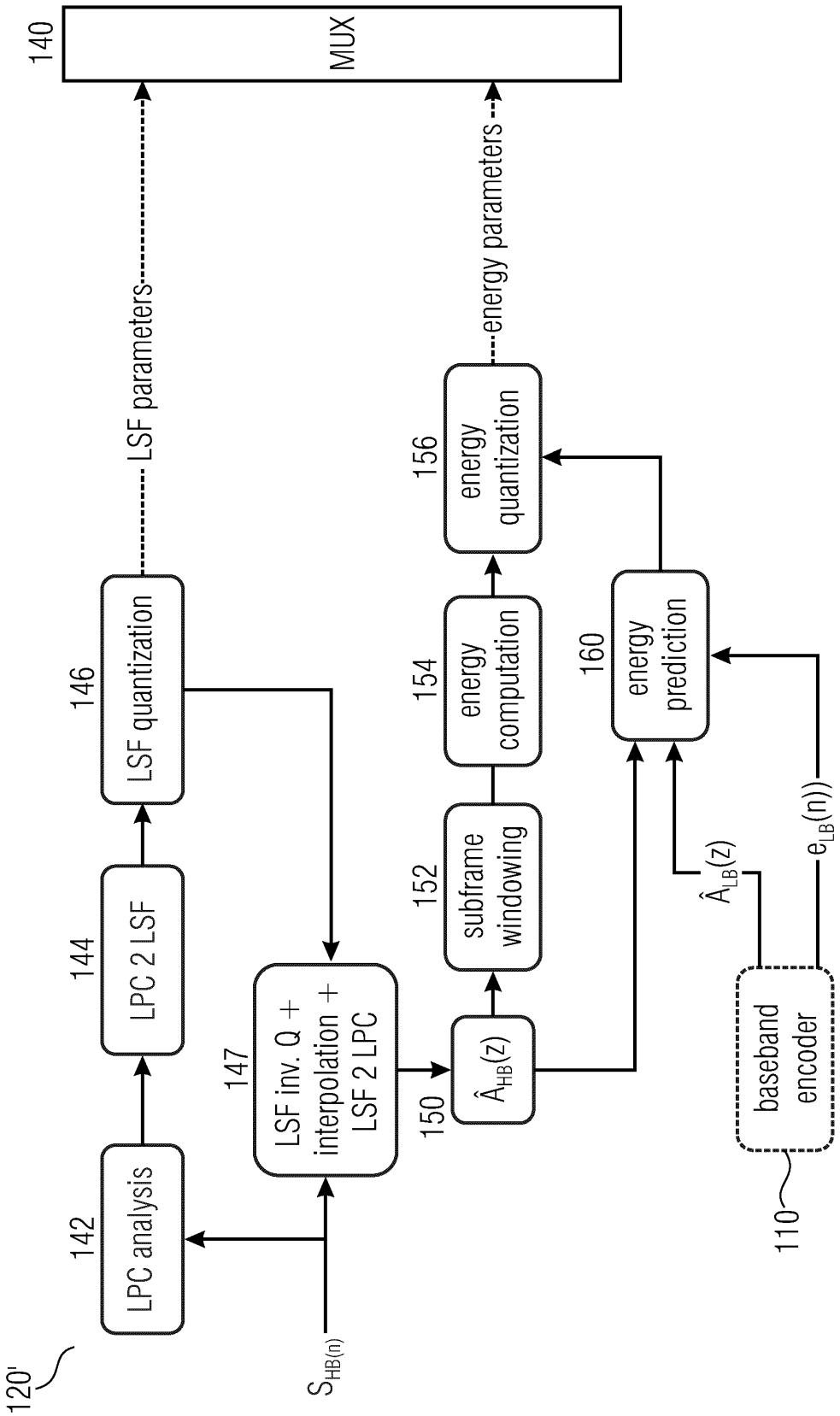

140
MUX
LSF parameters
energy parameters
146
LSF quantization
144
LPC 2 LSF
142
LPC analysis
147
LSF inv. Q + interpolation + LSF 2 LPC
150
$\hat{A}_{HB}(z)$
152
subframe windowing
154
energy computation
156
energy quantization
160
energy prediction
$\hat{A}_{LB}(z)$
$e_{LB}(n)$
baseband encoder
110
120'
$S_{HB(n)}$

Fig. 5

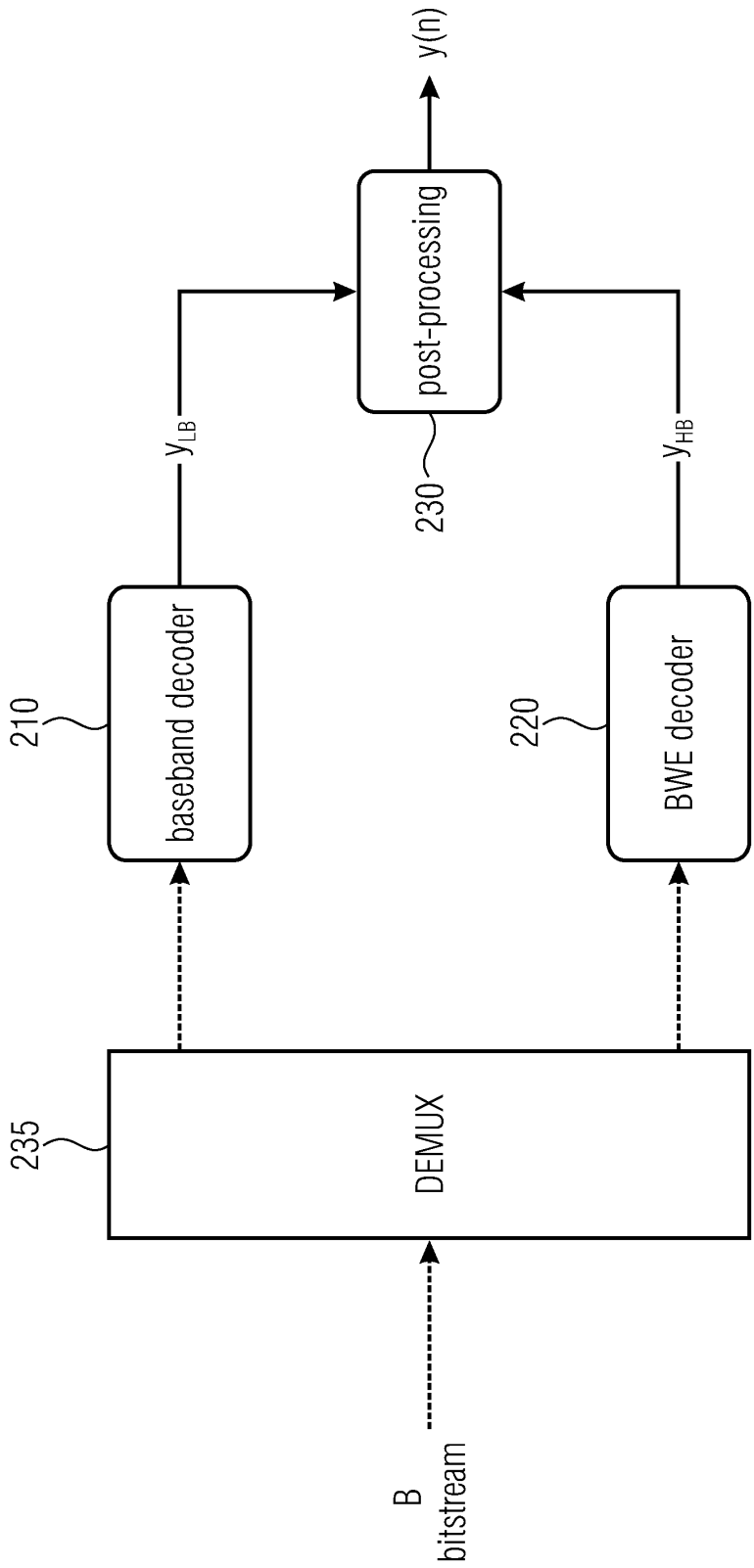

y(n)
post-processing
230
y_{LB}
y_{HB}
baseband decoder
BWE decoder
210
220
DEMUX
235
B
bitstream

Fig. 6

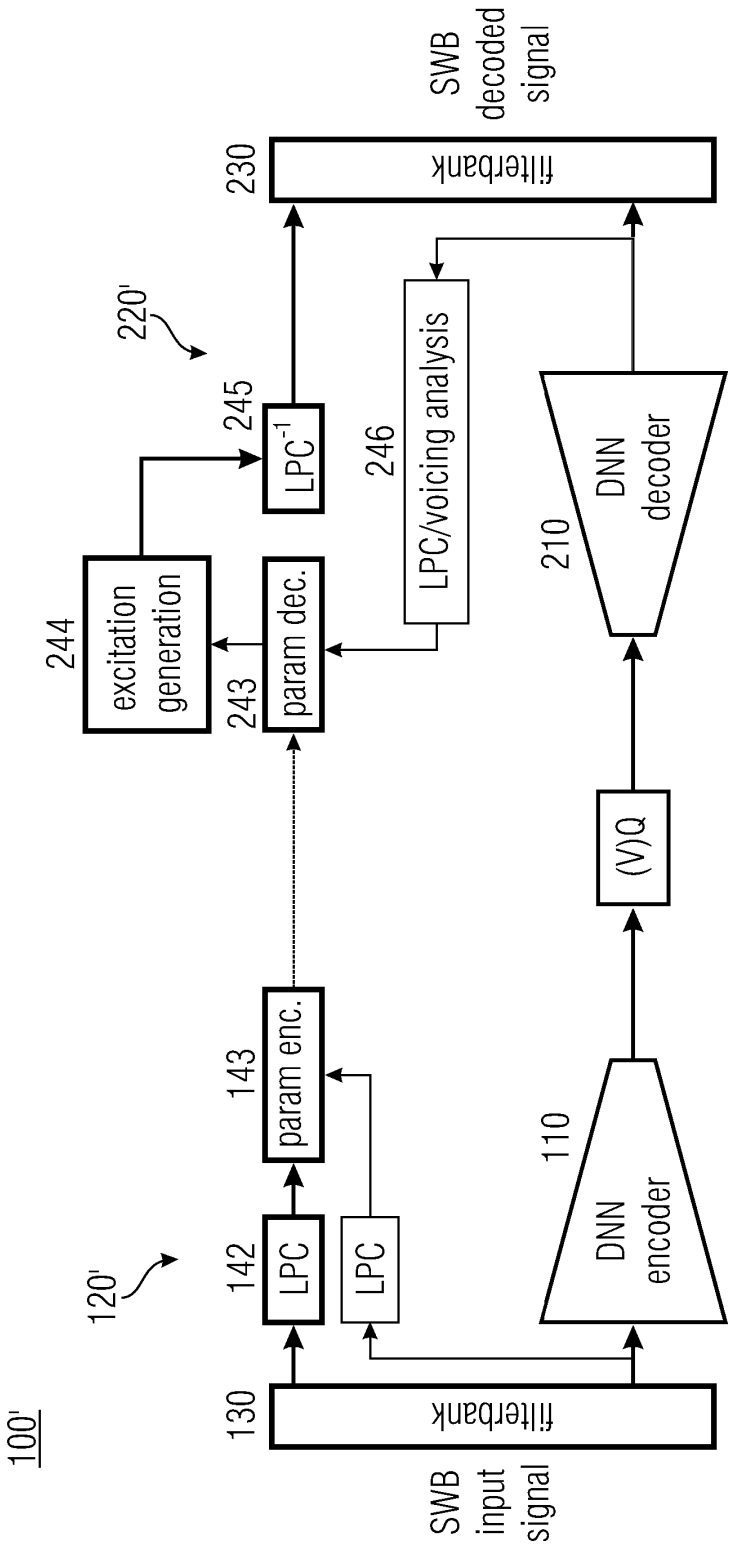

100'
120'
130
filterbank
SWB input signal
142
LPC
LPC
143
param enc.
110
DNN encoder
(V)Q
210
DNN decoder
200'
SWB decoded signal
230
filterbank
220'
245
LPC-1
246
LPC/voicing analysis
244
excitation generation
243
param dec.

Fig. 7