

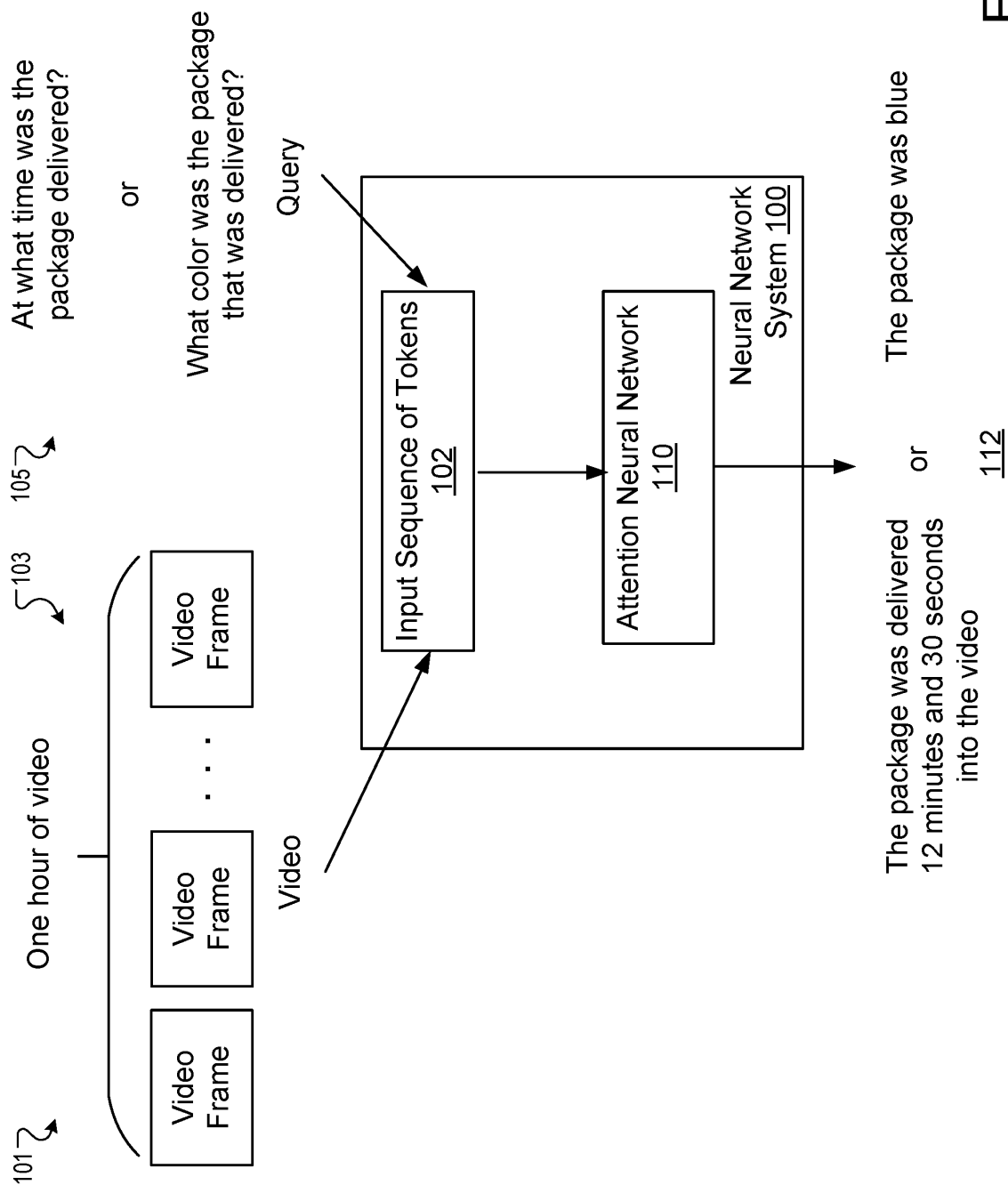


FIG. 1A

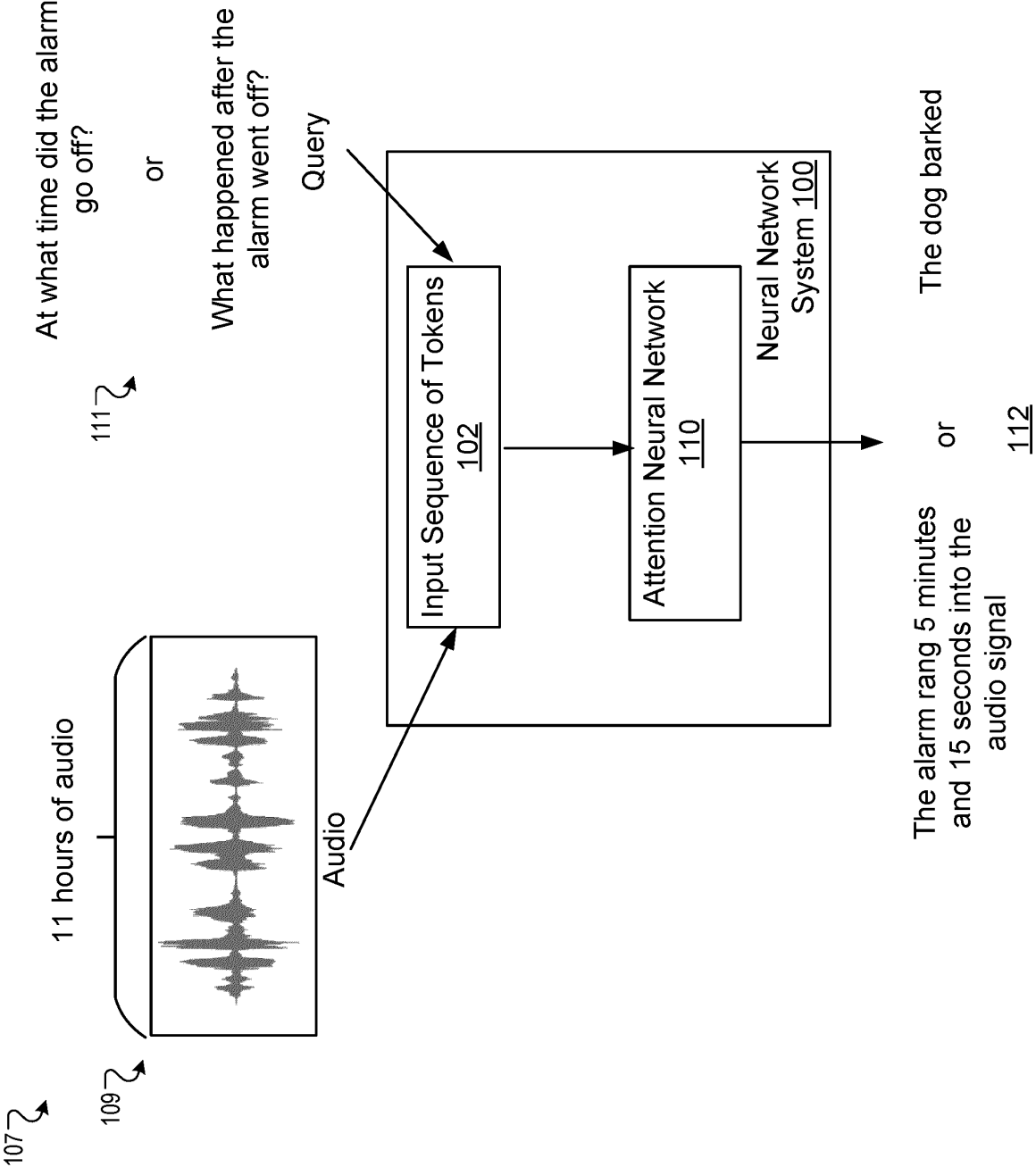


FIG. 1B

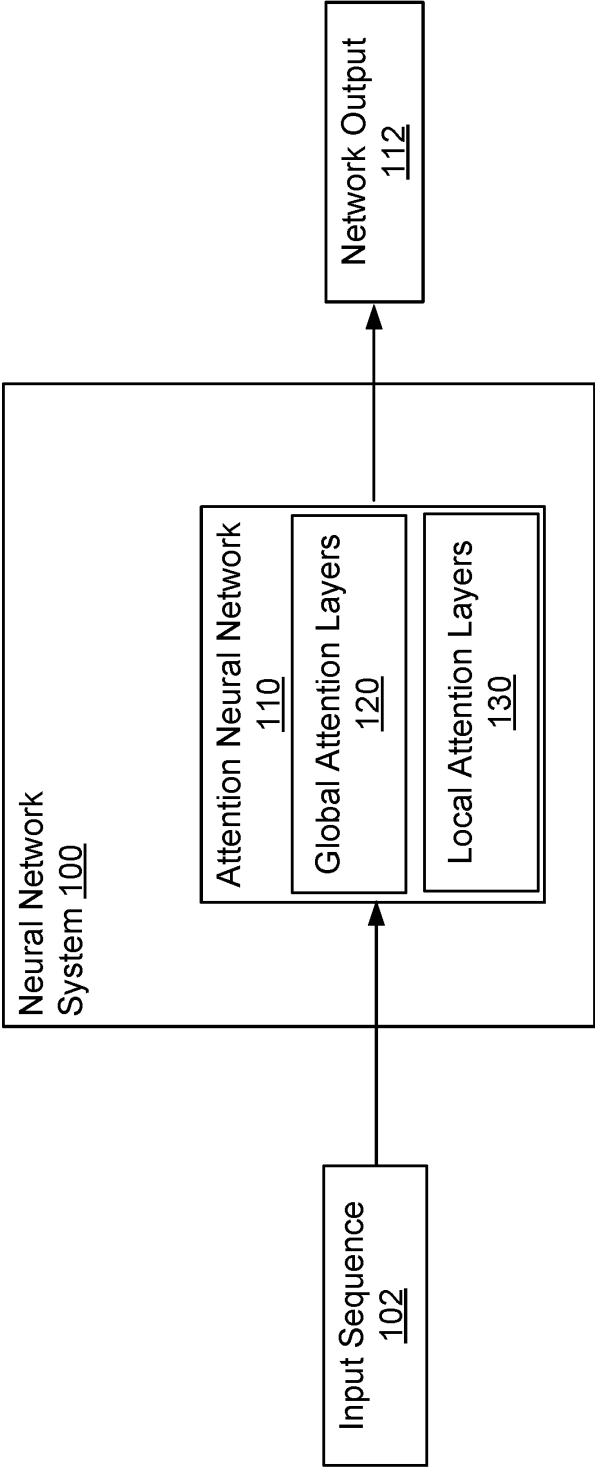


FIG. 1C

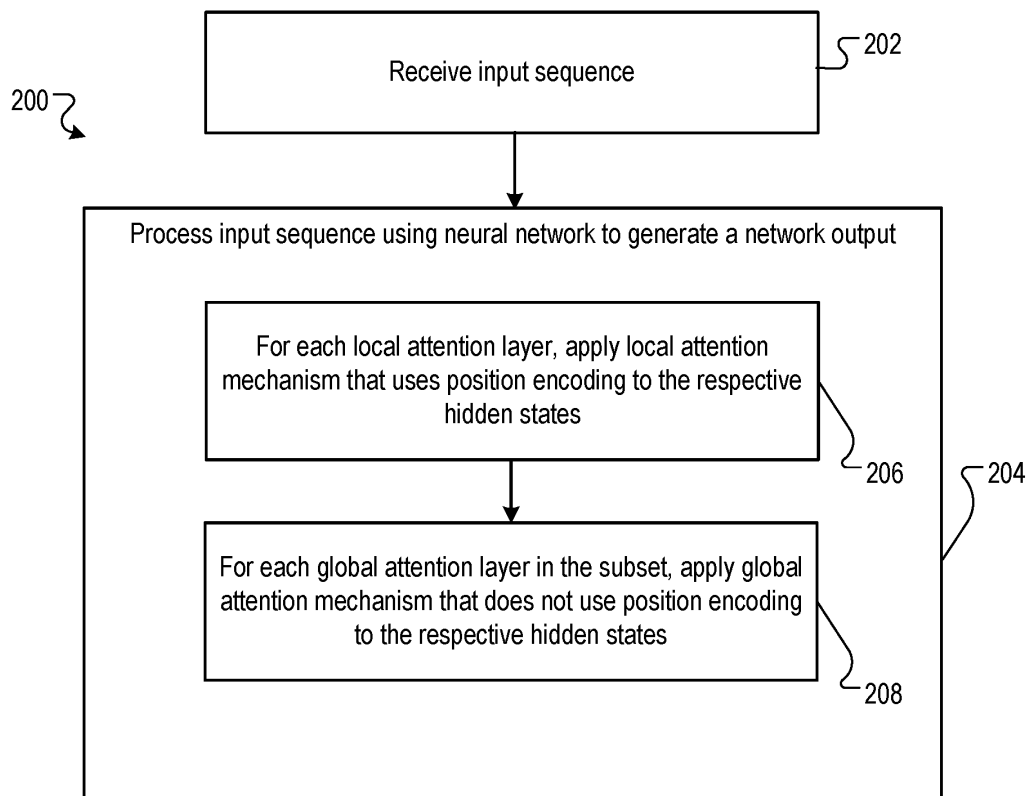


FIG. 2A

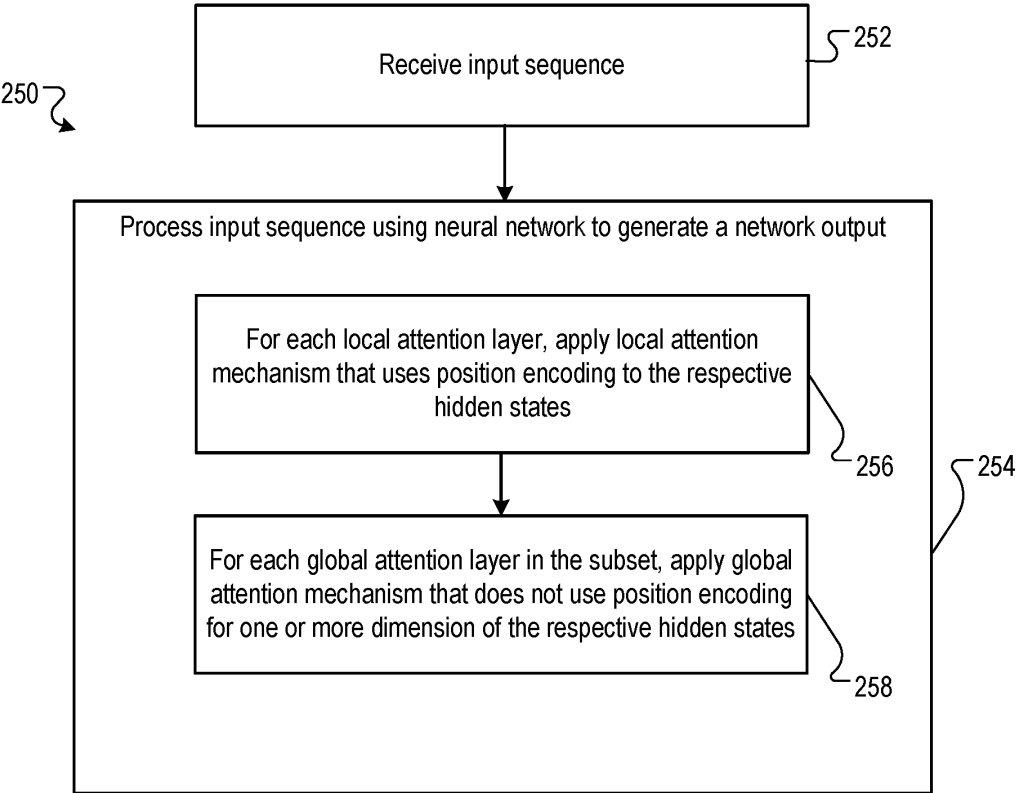


FIG. 2B

320 ↪

	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$
$\mathbf{x}_1$	$(\mathbf{W}_q \mathbf{x}_1)^T \mathbf{W}_k \mathbf{x}_1$	masked	masked
$\mathbf{x}_2$	$(\mathbf{W}_q \mathbf{x}_2)^T \mathbf{W}_k \mathbf{x}_1$	$(\mathbf{W}_q \mathbf{x}_2)^T \mathbf{W}_k \mathbf{x}_2$	masked
$\mathbf{x}_3$	$(\mathbf{W}_q \mathbf{x}_3)^T \mathbf{W}_k \mathbf{x}_1$	$(\mathbf{W}_q \mathbf{x}_2)^T \mathbf{W}_k \mathbf{x}_2$	$(\mathbf{W}_q \mathbf{x}_3)^T \mathbf{W}_k \mathbf{x}_3$

330 ↪

	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$
$\mathbf{x}_1$	$(\mathbf{R}_{\Theta,1}^d \mathbf{W}_q \mathbf{x}_1)^T (\mathbf{R}_{\Theta,1}^d \mathbf{W}_k \mathbf{x}_1)$	masked	masked
$\mathbf{x}_2$	$(\mathbf{R}_{\Theta,2}^d \mathbf{W}_q \mathbf{x}_2)^T (\mathbf{R}_{\Theta,1}^d \mathbf{W}_k \mathbf{x}_1)$	$(\mathbf{R}_{\Theta,2}^d \mathbf{W}_q \mathbf{x}_2)^T (\mathbf{R}_{\Theta,2}^d \mathbf{W}_k \mathbf{x}_2)$	masked
$\mathbf{x}_3$	masked	$(\mathbf{R}_{\Theta,3}^d \mathbf{W}_q \mathbf{x}_2)^T (\mathbf{R}_{\Theta,2}^d \mathbf{W}_k \mathbf{x}_2)$	$(\mathbf{R}_{\Theta,3}^d \mathbf{W}_q \mathbf{x}_3)^T (\mathbf{R}_{\Theta,3}^d \mathbf{W}_k \mathbf{x}_3)$

FIG. 3

400

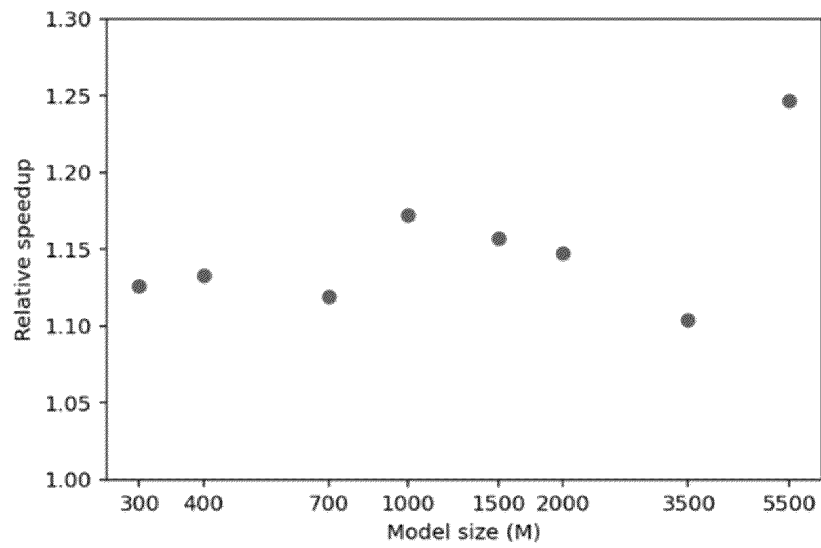


FIG. 4A

450 ↪

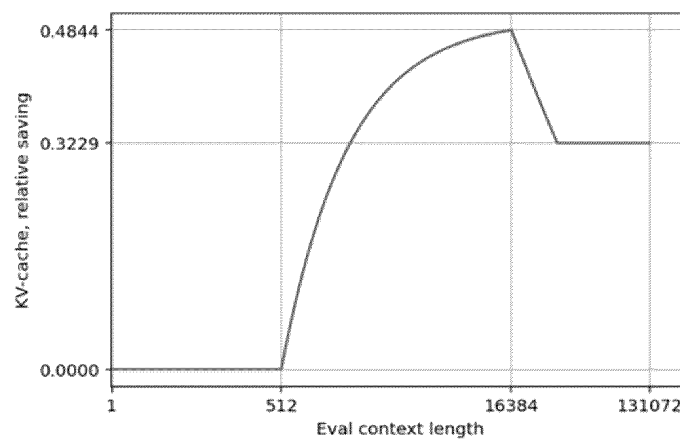


FIG. 4B

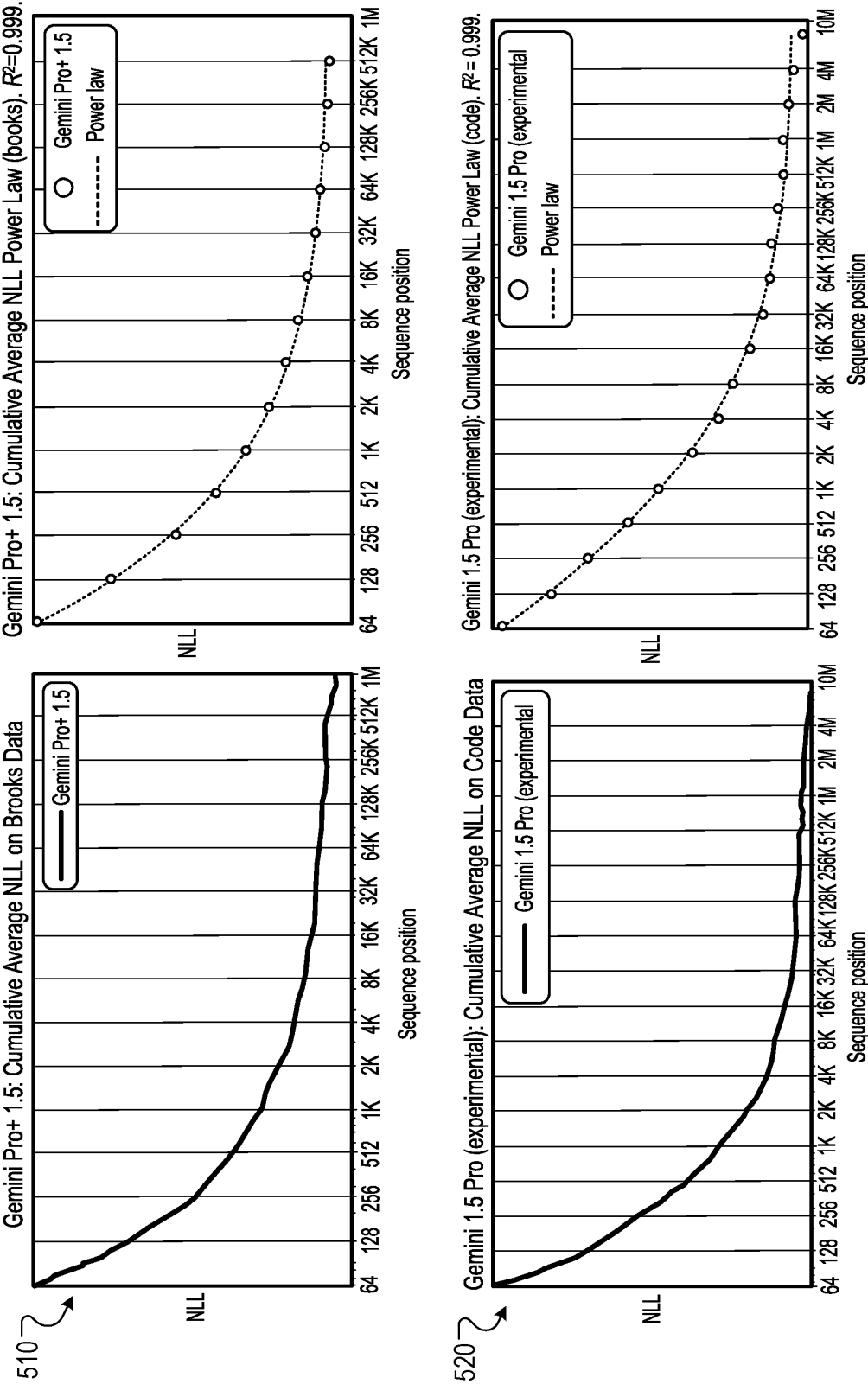


FIG. 5

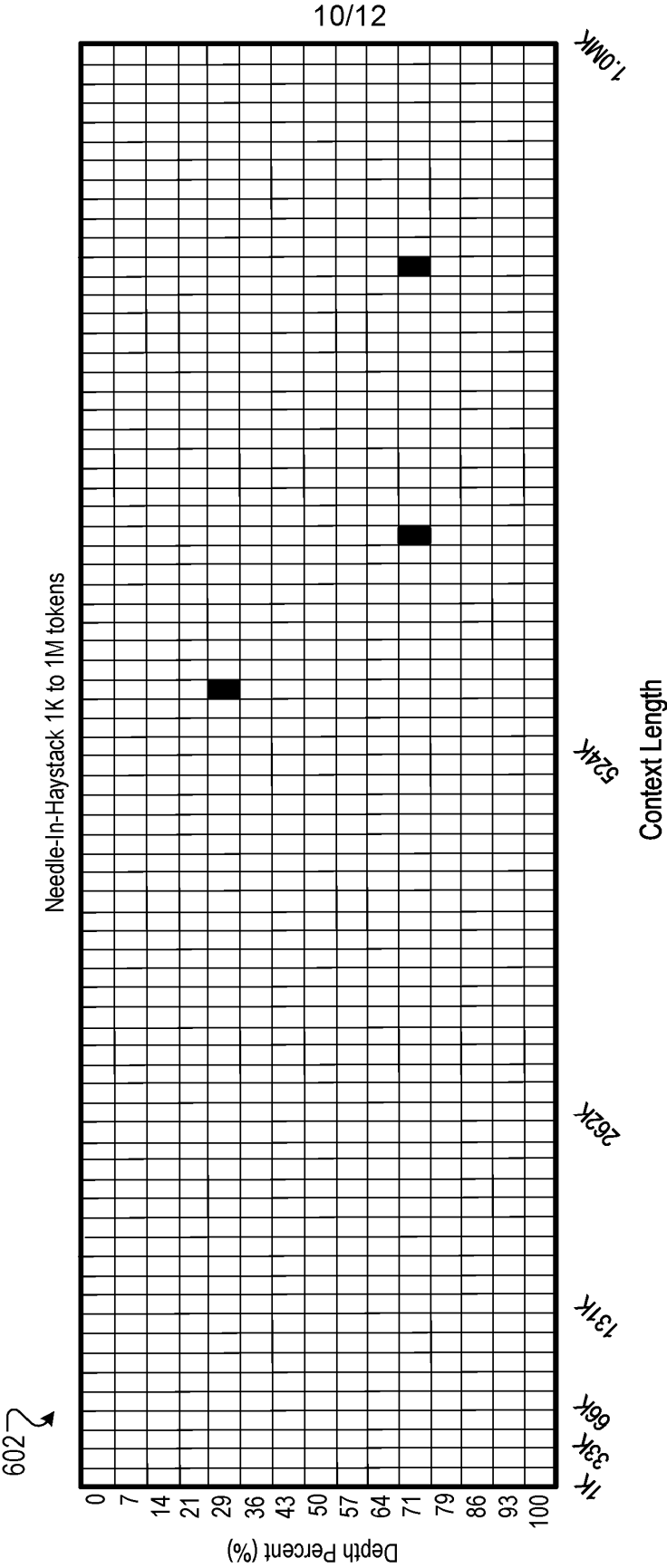


FIG. 6

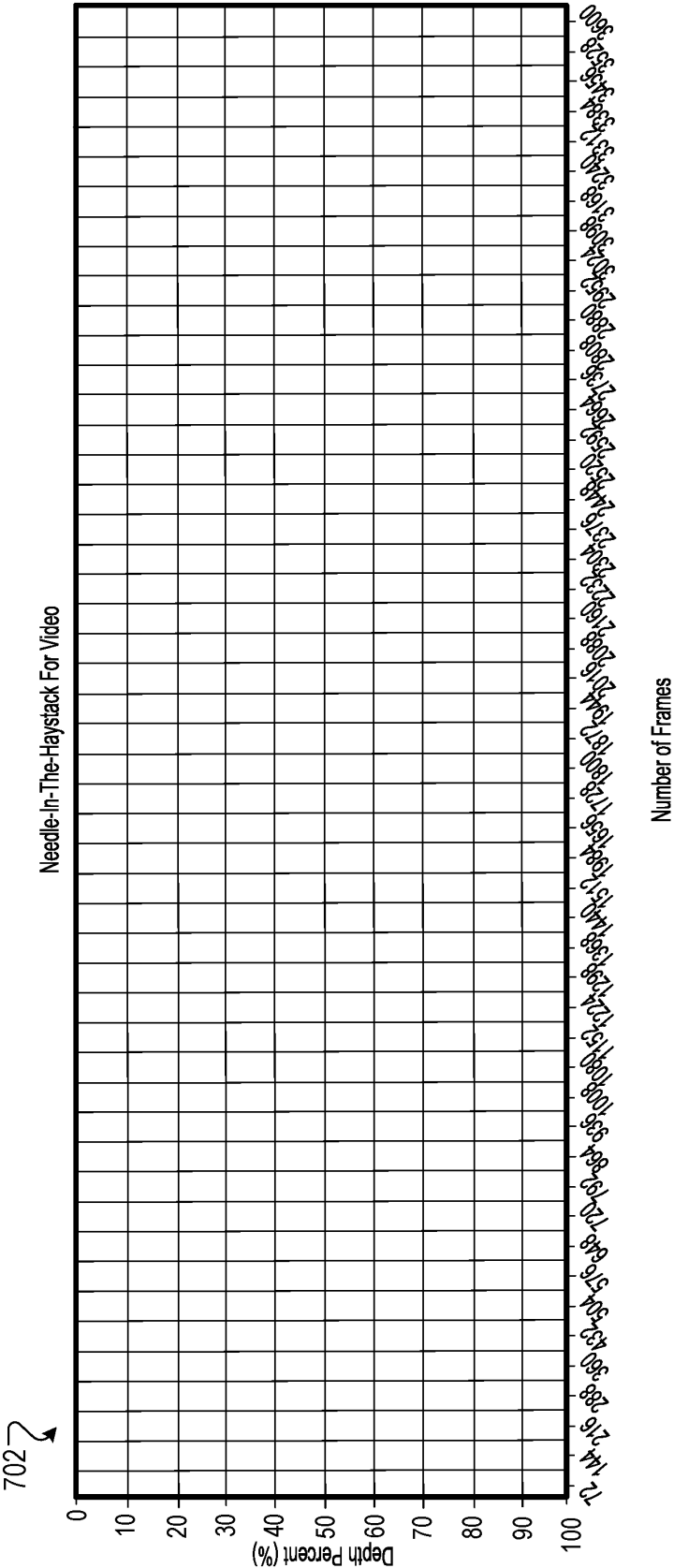


FIG. 7

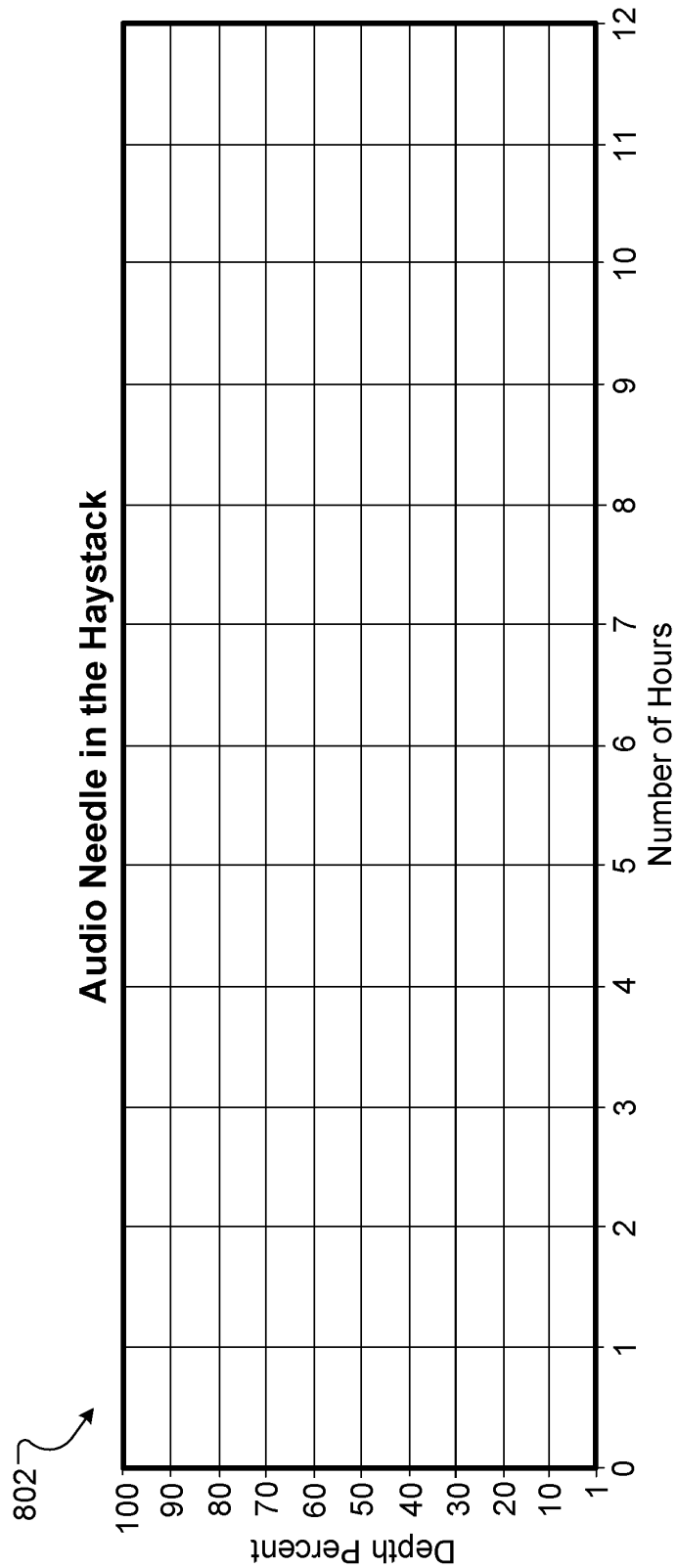


FIG. 8