

split attention probability for all attention heads

