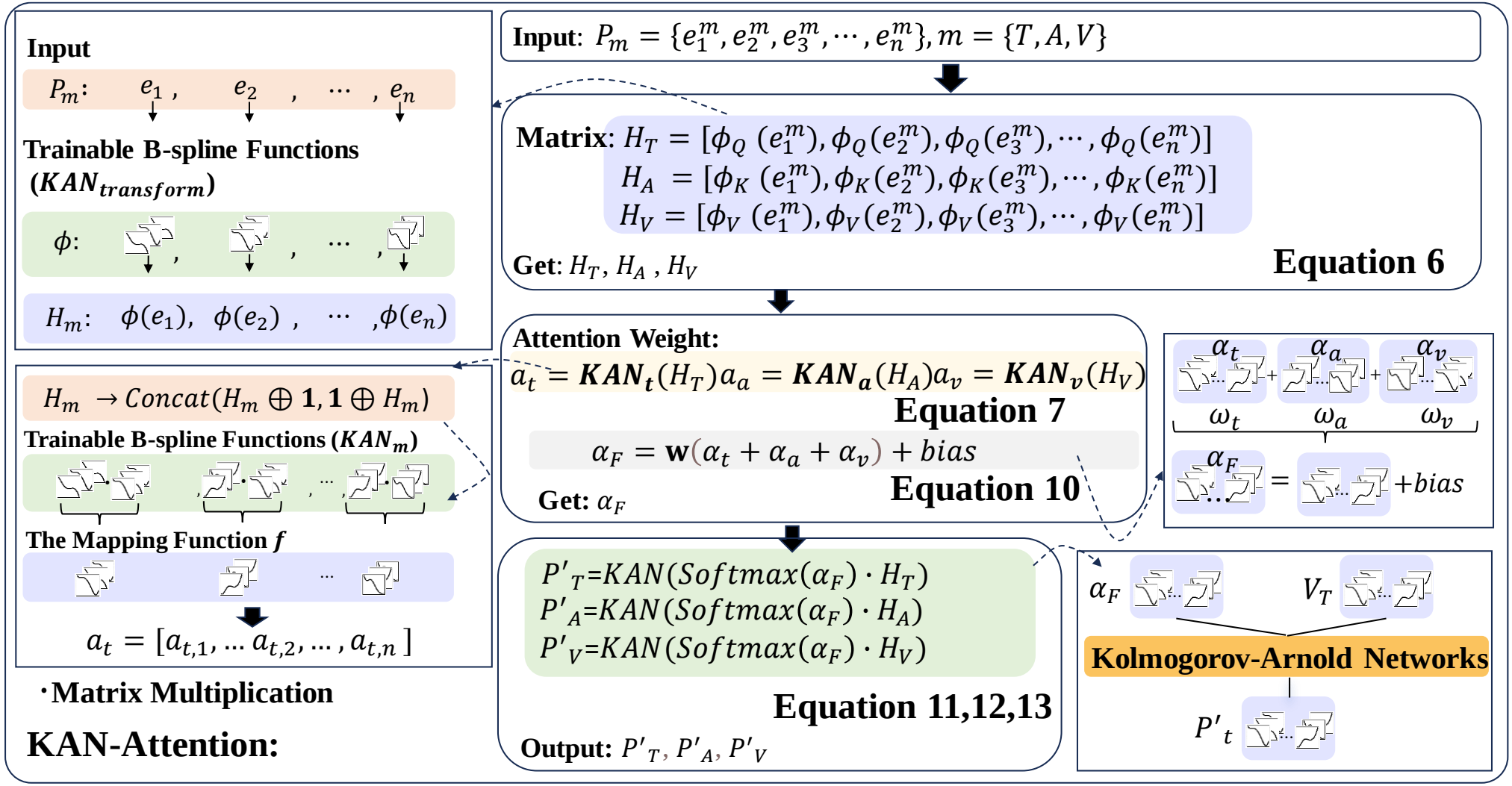
Input
$P_m$: $e_1$ , $e_2$ , … , $e_n$
Trainable B-spline Functions ($KAN_{transform}$)
$\phi$: , , … ,
$H_m$: $\phi(e_1)$, $\phi(e_2)$ , … ,$\phi(e_n)$
$H_m \rightarrow Concat(H_m \oplus \mathbf{1}, \mathbf{1} \oplus H_m)$
Trainable B-spline Functions ($KAN_m$)
The Mapping Function $f$
$a_t = [a_{t,1}, \dots a_{t,2}, \dots, a_{t,n}]$
·Matrix Multiplication
KAN-Attention:
Input: $P_m = \{e_1^m, e_2^m, e_3^m, \cdots, e_n^m\}, m = \{T, A, V\}$
Matrix: $H_T = [\phi_Q(e_1^m), \phi_Q(e_2^m), \phi_Q(e_3^m), \cdots, \phi_Q(e_n^m)]$
$H_A = [\phi_K(e_1^m), \phi_K(e_2^m), \phi_K(e_3^m), \cdots, \phi_K(e_n^m)]$
$H_V = [\phi_V(e_1^m), \phi_V(e_2^m), \phi_V(e_3^m), \cdots, \phi_V(e_n^m)]$
Get: $H_T$, $H_A$ , $H_V$
Equation 6
Attention Weight:
$a_t = \boldsymbol{KAN_t}(H_T)a_a = \boldsymbol{KAN_a}(H_A)a_v = \boldsymbol{KAN_v}(H_V)$
Equation 7
$\alpha_F = \mathbf{w}(\alpha_t + \alpha_a + \alpha_v) + bias$
Get: $\alpha_F$
Equation 10
$\alpha_t$ $\alpha_a$ $\alpha_v$
$\omega_t$ $\omega_a$ $\omega_v$
$\alpha_F$ = +bias
$P'_T = KAN(Softmax(\alpha_F) \cdot H_T)$
$P'_A = KAN(Softmax(\alpha_F) \cdot H_A)$
$P'_V = KAN(Softmax(\alpha_F) \cdot H_V)$
Equation 11,12,13
Output: $P'_T$, $P'_A$, $P'_V$
$\alpha_F$ $V_T$
Kolmogorov-Arnold Networks
$P'_t$