SUPPLEMENTARY MATERIALS: RNA-RNA INTERACTION PREDICTION: PARTITION FUNCTION AND BASE PAIR PAIRING PROBABILITIES

FENIX W.D. HUANG [1], JING QIN [1], CHRISTIAN M. REIDYS [1,2], AND PETER F. STADLER[3,4,5,6]

1. Preliminaries

1.1. **Energy model.** Let us review the energy model, implemented in `rip`. Its is an extension of the standard energy model of RNA secondary structures and recognizes the following loop-types:

(1) *Hairpin-loop*: a hairpin loop $\mathsf{Ha}_{i,j}$ has tabulated energies $G^{\mathsf{Ha}}_{i,j}$ depending on their sequence and length.
(2) *Interior-loop*: an interior loop $\mathsf{Int}_{i_1,j_1;i_2,j_2}$ also have tabulated energies $G^{\mathsf{Int}}_{i_1,j_1;i_2,j_2}$.
(3) *Multi-loop*: a multi-loop M_{i_0,j_0} has energy $\alpha_1 + \alpha_2(t+1) + \alpha_3 c_2$, where $t = |E^i_{R[i_0,j_0]}|$ ("branching order") inside $R[i_0, j_0]$ and c_2 is the number of isolated vertices contained in $R[i_0, j_0]$.
(4) *Kissing-loop*: a kissing-loop K_{i_0,j_0} has energy $\beta_1+\beta_2(t+1)+\beta_3 c_2$, where $t = |E^i_{R[i_0,j_0]}|$ and c_2 is the number of isolated vertices contained in $R[i_0, j_0]$, analogous to the parametrization of multiloops.
(5) *Hybrid*: a hybrid $\mathsf{Hy}_{i_1,i_\ell;j_1,j_\ell}$ has energy $G^{\mathsf{Hy}}_{i_1,i_\ell;j_1,j_\ell} = \sigma_0 + \sigma \sum_\theta G^{\mathsf{Int}}_{i_\theta,i_{\theta+1};j_\theta,j_{\theta+1}}$, where a intermolecular interior loop formed by $R_{i_\theta} S_{j_\theta}$ and $R_{i_{\theta+1}} S_{j_{\theta+1}}$ is treated like an interior loop $\mathsf{Int}_{i_\theta,j_\theta;i_{\theta+1},j_{\theta+1}}$ with an affine scaling σ.

1.2. **Structural components.** In Figure 1 we display the twelve basic structural components: **A**, **B**: maximal secondary structure segments, $R[i, j]$ and $S[r, s]$, respectively; **C**: an arbitrary joint structure $J_{i,j;r,s}$; **D**: a right-tight structures $J^{RT}_{i,j;r,s}$; **E**: a double-tight structure $J^{DT}_{i,j;r,s}$; **F**: a tight structure having type $\bigtriangledown$, $\bigtriangleup$ or $\square$, respectively; **G**: a tight structure, $J^{\square}_{i,j;r,s}$, of type $\square$; **H**: a tight

Date: April, 2009.

structure, $J^{\bigtriangledown}_{i,j;r,s}$, of type $\bigtriangledown$; **J**: a tight structure, $J^{\bigtriangleup}_{i,j;r,s}$, of type $\bigtriangleup$; **K**: exterior arc; **L**: isolated segment; **M**: pair of secondary segments, one of which containing at least one arc.

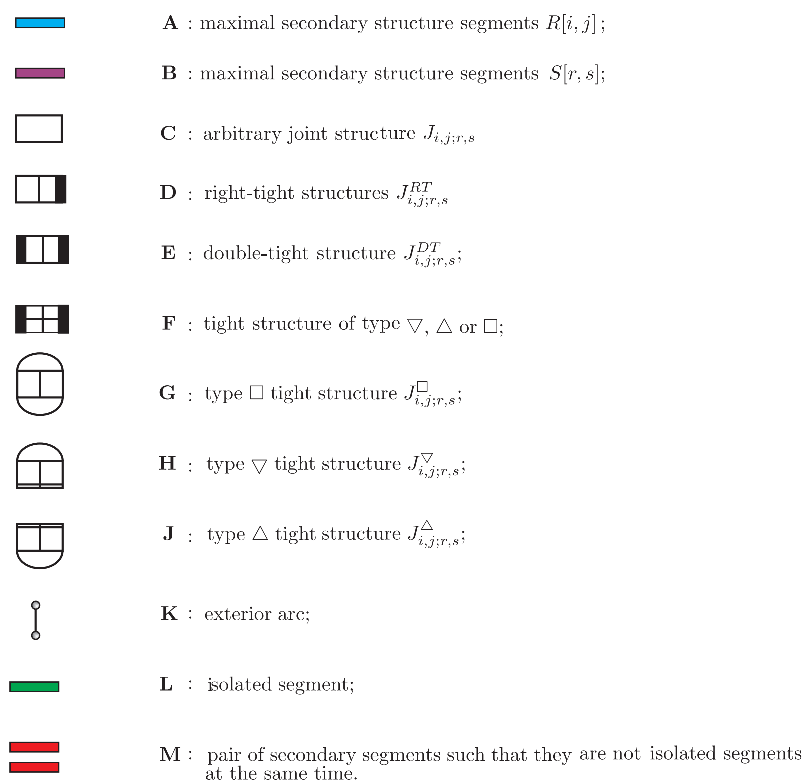

FIGURE 1. The panel displays the twelve basic types of structural components.

2. Recurrences

The complete set of 4D-storage arrays and 2D-storage array for the partition function are displayed in the Tables 1-5.

TABLE 1. Tight structures, $Q^{T}_{i,j;r,s}$: 15 4D-arrays, where $T \in \{\bigtriangledown, \bigtriangleup, \Box\}$.

$Q^{T,EE}$	$Q^{T,ME}$	$Q^{T,EM}$	$Q^{T,FE}$	$Q^{T,EF}$
$Q^{T,MM}$	$Q^{T,MF}$	$Q^{T,FM}$	$Q^{T,FF}$	$Q^{T,EK}$
$Q^{T,MK}$	$Q^{T,FK}$	$Q^{T,KE}$	$Q^{T,KM}$	$Q^{T,KF}$

TABLE 2. Right-tight joint structures, $Q^{RT}_{i,j;r,s}$: 24 4D-arrays.

$Q^{RT,EE}$	$Q^{RT,ME}$	$Q^{RT,EM}$	$Q^{RT,FE}$	$Q^{RT,EF}$	$Q^{RT,MM}$
$Q^{RT,MF}$	$Q^{RT,FM}$	$Q^{RT,FF}$	$Q^{RT,EK}$	$Q^{RT,MK}$	$Q^{RT,FK}$
$Q^{RT,KE}$	$Q^{RT,KM}$	$Q^{RT,KF}$	$Q^{RT,KK}$	$Q^{RT,EEA}$	$Q^{RT,EEB}$
$Q^{RT,EKA}$	$Q^{RT,EKB}$	$Q^{RT,KEA}$	$Q^{RT,KEB}$	$Q^{RT,KKA}$	$Q^{RT,KKB}$

TABLE 3. Double-tight joint structures, $Q^{DT}_{i,j;r,s}$: 18 4D-matrices.

$Q^{DT,ME}$	$Q^{DT,EM}$	$Q^{DT,MM}$	$Q^{DT,MF}$	$Q^{DT,FM}$	$Q^{DT,EK}$
$Q^{DT,MK}$	$Q^{DT,FK}$	$Q^{DT,KE}$	$Q^{DT,KM}$	$Q^{DT,KF}$	$Q^{DT,KK}$
$Q^{DT,EKA}$	$Q^{DT,EKB}$	$Q^{DT,KEA}$	$Q^{DT,KEB}$	$Q^{DT,KKA}$	$Q^{DT,KKB}$

TABLE 4. Joint structures, $Q^{I}_{i,j;r,s}$: 16 4D-arrays.

$Q^{I,EE}$	$Q^{I,ME}$	$Q^{I,EM}$	$Q^{I,FE}$	$Q^{I,EF}$	$Q^{I,MM}$	$Q^{I,MF}$	$Q^{I,FM}$
$Q^{I,FF}$	$Q^{I,EK}$	$Q^{I,MK}$	$Q^{I,FK}$	$Q^{I,KE}$	$Q^{I,KM}$	$Q^{I,KF}$	$Q^{I,KK}$

The complete set of recursions comprises for tight structures $Q^{T}_{i,j;r,s}$, 15 4D-arrays, for right-tight joint structures $Q^{RT}_{i,j;r,s}$, 24 4D-arrays, for double-tight structures $Q^{DT}_{i,j;r,s}$, 18 4D-arrays, 16 4D-arrays for arbitrary interaction structures $Q^{I}_{i,j;r,s}$ and 8 2D-arrays for secondary segments.

Structure-type	recurrence-formula (symbolic)
$J^{\bigtriangledown}_{i,j;h,\ell}$	Figure 3
$J^{\bigtriangleup}_{i,j;h,\ell}$	Figure 4
$J^{\Box}_{i,j;h,\ell}$	Figure 5
$J^{DT}_{i,j;h,\ell}$	Figure 6
$J^{RT}_{i,j;h,\ell}$	Figure 7
$J_{i,j;h,\ell}$	Figure 8

Table 5. Secondary segments: 8 2D-arrays.

Q^R	$Q^{R,b}$	$Q^{R,M}$	$Q^{R,F}$
Q^S	$Q^{S,b}$	$Q^{S,M}$	$Q^{S,F}$

3. Computation of the base pairing probabilities

In contrast to the computation of the partition function "from the inside to the outside", the computation of the base pairing probabilities (bpp) is obtained "from the outside to the inside". Let $\mathbb{J}_{i,j;h,\ell}^{\xi,Y_1Y_2Y_3}$ be the set of substructures $J_{i,j;h,\ell} \subset J_{1,N;1,M}$ such that $J_{i,j;h,\ell}$ appears in $T_{1,N;1,M}$ as an interaction structure of type $\xi \in \{DT, RT, \triangledown, \triangle, \square, \circ\}$ with loop-subtypes $Y_1, Y_2 \in \{\mathsf{M}, \mathsf{K}, \mathsf{F}\}$ on the sub-intervals $R[i,j]$ and $S[h,\ell]$, $Y_3 \in \{\mathsf{A}, \mathsf{B}\}$. Let $\mathbb{P}_{i,j;h,\ell}^{\xi,Y_1Y_2Y_3}$ be the probability of $\mathbb{J}_{i,j;h,\ell}^{\xi,Y_1Y_2Y_3}$. For instance, $\mathbb{P}_{i,j;h,\ell}^{RT,\mathsf{MKA}}$ is the sum over all the probabilities of substructures $J_{i,j;h,\ell} \in T_{1,N;1,M}$ such that $J_{i,j;h,\ell}$ is a right-tight structure of type $r\mathsf{A}$ and $R[i,j]$, $S[h,\ell]$ are enclosed by a multi-loop and kissing loop, respectively.

Algorithm 1 constructs recursively all 4D-arrays $\mathrm{P}_{i,i+j;r,r+s}^{\xi,Y_1Y_2Y_3}$. This is obtained via the corresponding arrays of partition functions over the respective subcomplexes and the quantities $\mathrm{P}_{i,i+j;r,r+s}^{\xi,Y_1Y_2Y_3}$ from the outisde to the inside. In other words Algorithm 1 facilitates the recursive translation of the 4D-arrays of partition functions into base pairing probabilities. By construction we have

$$\mathbb{P}_{i,i+j;r,r+s}^{\xi,Y_1Y_2Y_3} = \mathrm{P}_{i,i+j;r,r+s}^{\xi,Y_1Y_2Y_3}. \tag{3.1}$$

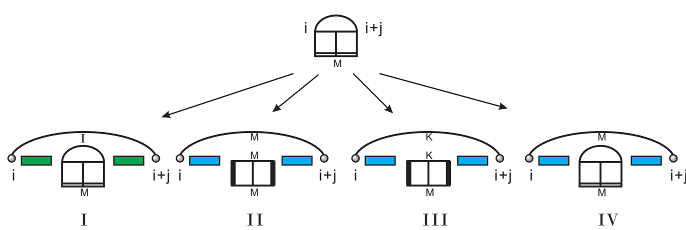

Figure 2. Further refinement: the four decompositions of $J_{i,j;r,s}^{\triangledown,\mathsf{M}}$ via Procedure (b). These cases correspond to the four contributions in Algorithm 1).

Algorithm 1 Case I to Case IV correspond to the fours cases showed in Figure 2.

$j \leftarrow$ lengthR-1
while $j \geq 0$ **do**
  for $i \leftarrow 1$ to lengthR $- j$ **do**
    $s \leftarrow$ lengthS-1
    while $s \geq 0$ **do**
      for $r \leftarrow 1$ to lengthS $- s$ **do**
        ⋮
        if $Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \neq 0$ **then**
          for $h \leftarrow i+1$ to $i+j-1$ **do**
            for $\ell \leftarrow h$ to $i+j-1$ **do**
              $Q \leftarrow Q^{\triangledown,\mathsf{M}}_{h,\ell;r,r+s} \cdot e^{-G^{\mathtt{Int}}_{i,i+j;h,\ell}}$
              $\mathrm{P}^{\triangledown,\mathsf{M}}_{h,\ell;r,r+s} \leftarrow \mathrm{P}^{\triangledown,\mathsf{M}}_{h,\ell;r,r+s} + \mathrm{P}^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \cdot Q/Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s}$ {Case I}
              $Q \leftarrow Q^{\mathsf{R,M}}_{i+1,h-1} \cdot Q^{\mathsf{R,M}}_{\ell+1,i+j-1} \cdot Q^{\triangledown,\mathsf{M}}_{h,\ell;r,r+s} \cdot \exp(-(\alpha_1 + 2\alpha_2)/RT)$
              $\mathrm{P}^{\triangledown,\mathsf{M}}_{h,\ell;r,r+s} \leftarrow \mathrm{P}^{\triangledown,\mathsf{M}}_{h,\ell;r,r+s} + \mathrm{P}^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \cdot Q/Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s}$
              $\mathrm{P}^{\mathsf{R,M}}_{i+1,h-1} \leftarrow \mathrm{P}^{\mathsf{R,M}}_{i+1,h-1} + \mathrm{P}^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \cdot Q/Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s}$
              $\mathrm{P}^{\mathsf{R,M}}_{\ell+1,i+j-1} \leftarrow \mathrm{P}^{\mathsf{R,M}}_{\ell+1,i+j-1} + \mathrm{P}^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \cdot Q/Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s}$ {Case II}
              $Q \leftarrow Q^{\mathsf{R,M}}_{i+1,h-1} \cdot Q^{\mathsf{R,M}}_{\ell+1,i+j-1} \cdot Q^{\mathsf{DT,MM}}_{h,\ell;r,r+s} \cdot \exp(-(\alpha_1 + \alpha_2)/RT)$
              $\mathrm{P}^{\mathsf{DT,MM}}_{h,\ell;r,r+s} \leftarrow \mathrm{P}^{\mathsf{DT,MM}}_{h,\ell;r,r+s} + \mathrm{P}^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \cdot Q/Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s}$
              $\mathrm{P}^{\mathsf{R,M}}_{i+1,h-1} \leftarrow \mathrm{P}^{\mathsf{R,M}}_{i+1,h-1} + \mathrm{P}^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \cdot Q/Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s}$
              $\mathrm{P}^{\mathsf{R,M}}_{\ell+1,i+j-1} \leftarrow \mathrm{P}^{\mathsf{R,M}}_{\ell+1,i+j-1} + \mathrm{P}^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \cdot Q/Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s}$ {Case III}
              $Q \leftarrow Q^{\mathsf{R,F}}_{i+1,h-1} \cdot Q^{\mathsf{R,F}}_{\ell+1,i+j-1} \cdot Q^{\mathsf{DT,KM}}_{h,\ell;r,r+s} \cdot \exp(-(\beta_1 + \beta_2)/RT)$
              $\mathrm{P}^{\mathsf{DT,KM}}_{h,\ell;r,r+s} \leftarrow \mathrm{P}^{\mathsf{DT,KM}}_{h,\ell;r,r+s} + \mathrm{P}^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \cdot Q/Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s}$
              $\mathrm{P}^{\mathsf{R,F}}_{i+1,h-1} \leftarrow \mathrm{P}^{\mathsf{R,F}}_{i+1,h-1} + \mathrm{P}^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \cdot Q/Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s}$
              $\mathrm{P}^{\mathsf{R,F}}_{\ell+1,i+j-1} \leftarrow \mathrm{P}^{\mathsf{R,F}}_{\ell+1,i+j-1} + \mathrm{P}^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s} \cdot Q/Q^{\triangledown,\mathsf{M}}_{i,i+j;r,r+s}$ {Case IV}
            end for
          end for
        end if
        ⋮
      end for
      $s \leftarrow s - 1$
    end while
  end for
  $j \leftarrow j - 1$
end while

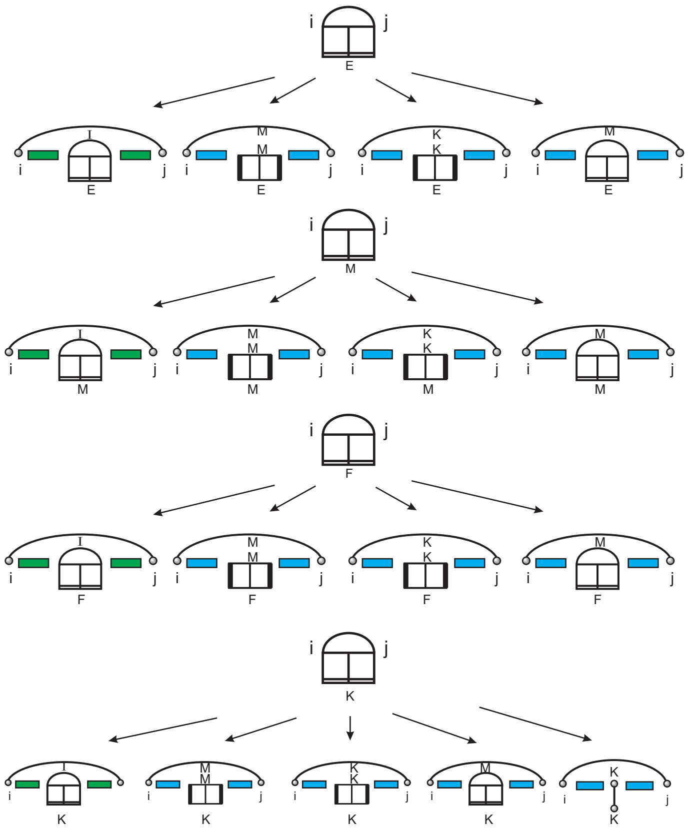

FIGURE 3. Decomposition for $J^{\triangledown}_{i,j;h,\ell}$.

References

[1]Center for Combinatorics, LPMC-TJKLC, Nankai University Tianjin 300071, P.R. China, [2]College of Life Science, Nankai University Tianjin 300071, P.R. China, [3]Bioinformatics Group, Department of Computer Science, and Interdisciplinary Center for Bioinformatics, University of Leipzig, Härtelstrasse 16-18, D-04107 Leipzig, Germany., [4]RNomics Group, Fraunhofer Institut for Cell Therapy and Immunology, Perlickstrasse 1,D-04103 Leipzig, Germany, [5]Inst. f. Theoretical Chemistry, University of Vienna, Währingerstrasse 17, A-1090 Vienna, Austria, [6]The Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, New Mexico

E-mail address: `duck@santafe.edu`

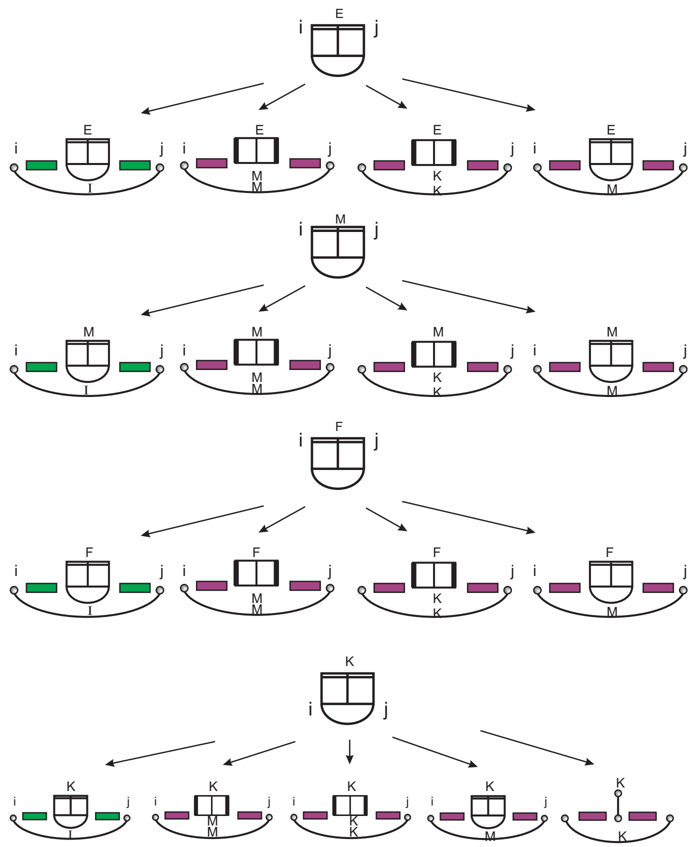

FIGURE 4. Decomposition for $J^{\triangle}_{i,j;h,\ell}$.

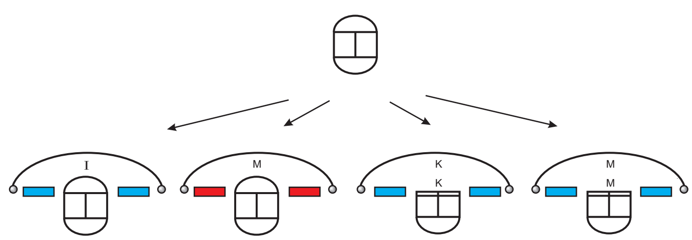

FIGURE 5. Decomposition for $J^{\square}_{i,j;h,\ell}$.

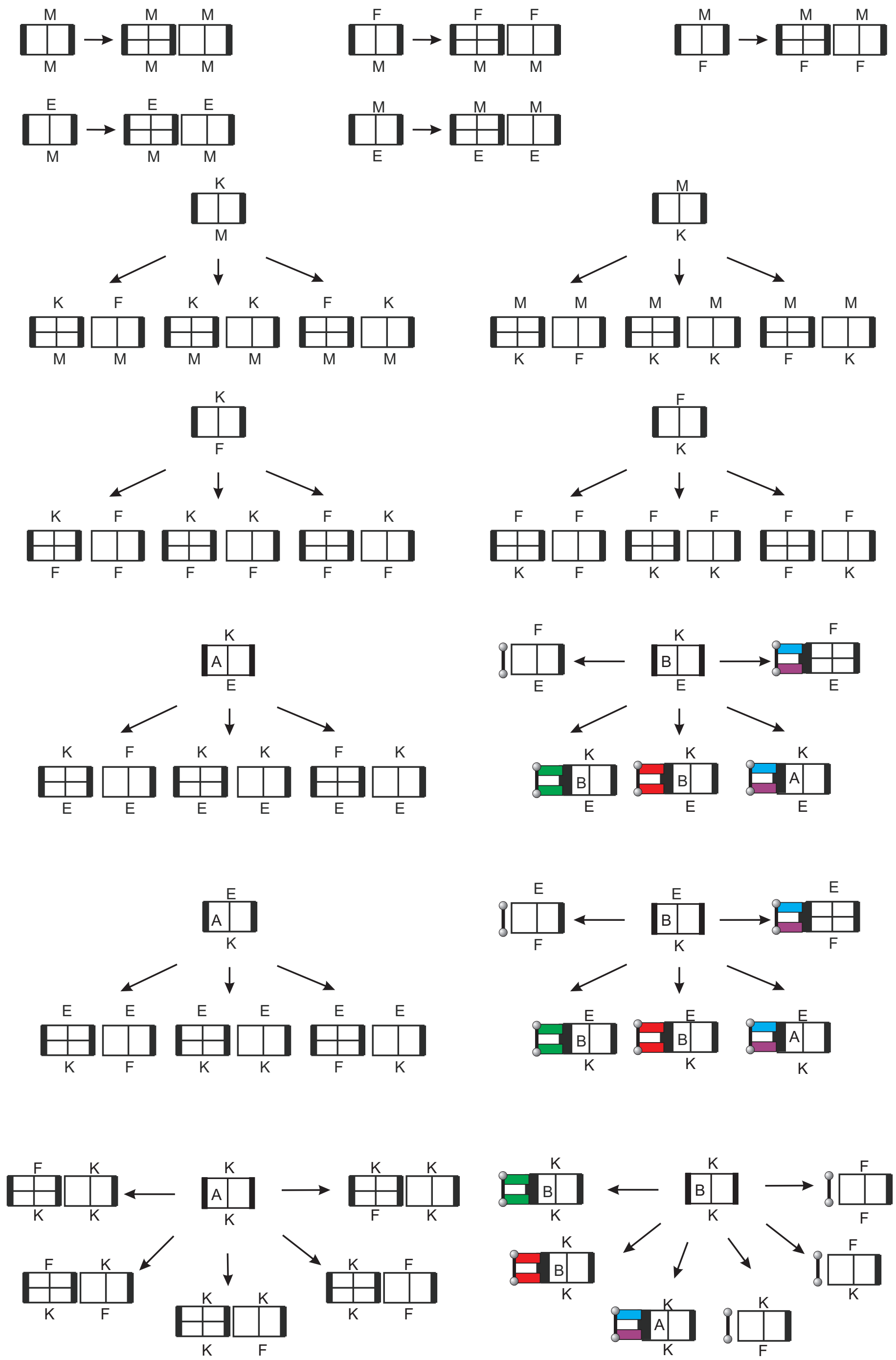

FIGURE 6. Decomposition for $J^{DT}_{i,j;h,\ell}$.

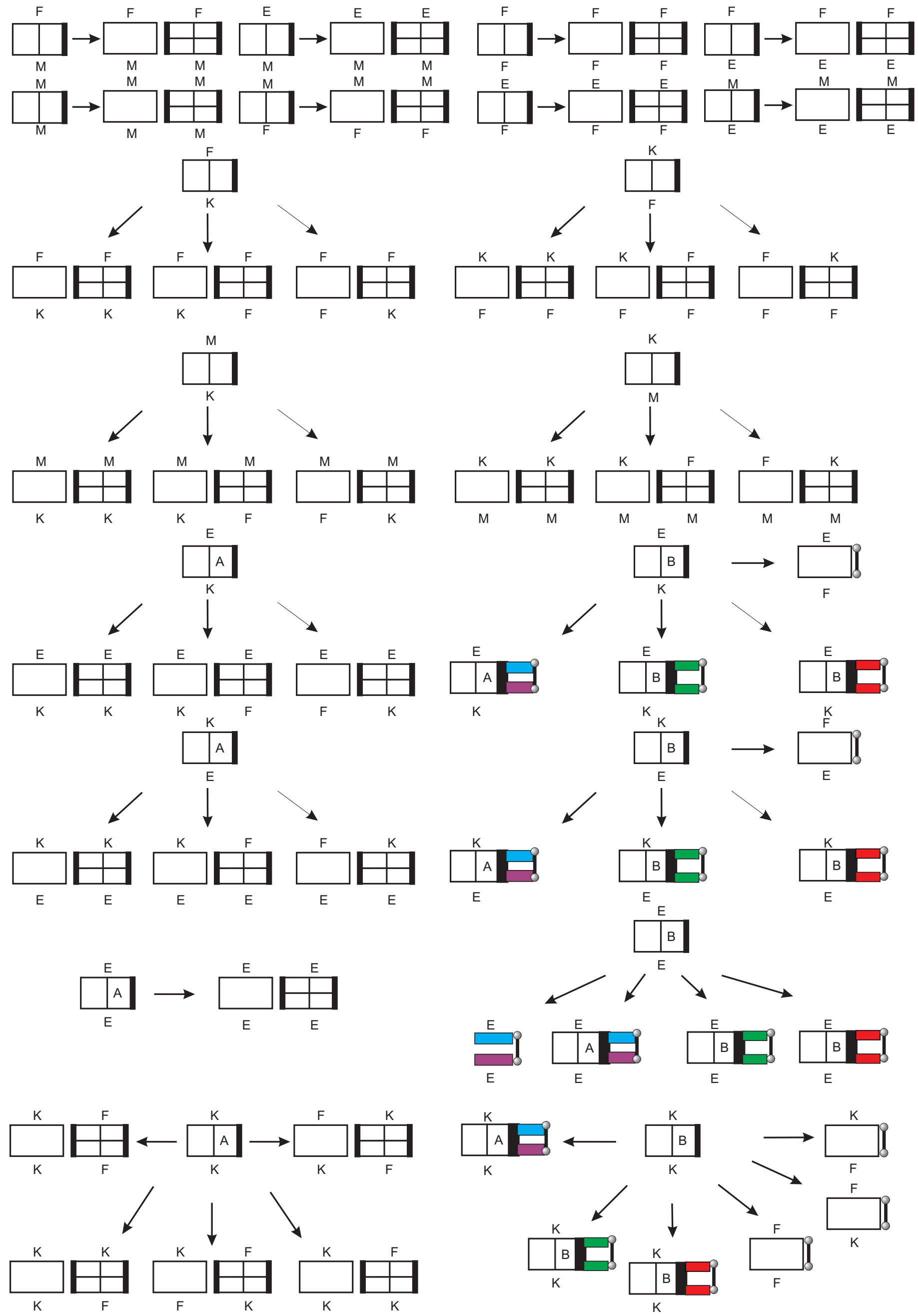

FIGURE 7. Decomposition for $J^{RT}_{i,j;h,\ell}$.

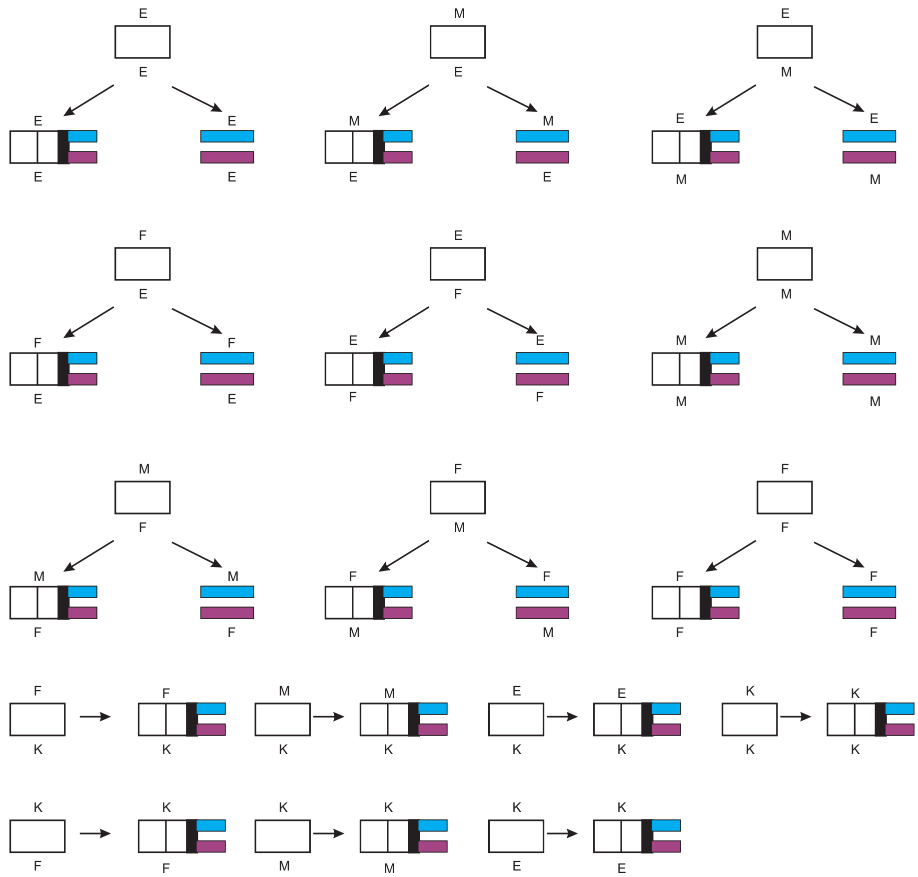

FIGURE 8. Decomposition for $J_{i,j;h,\ell}$.