Analysis of spectroscopic networks

Roland Tóbiás

MTA-ELTE Complex Chemical Systems Research Group

Related studies

(a) T. **Furtenbacher**, A. G. **Császár**, and J. Tennyson, *J. Mol. Spectrosc.* **2007**, *245*, 115-125.

(b) A. G. **Császár** and T. **Furtenbacher**, *J. Mol. Spectrosc.* **2011**, *266*, 99-103.

(c) T. **Furtenbacher** and A. G. **Császár**, *J. Quant. Spectr. Rad. Transfer*, **2012**, *113*, 929-935.

(d) T. **Furtenbacher** and A. G. **Császár**, *J. Mol. Struct.* **2012**, *1009*, 123-129.

(e) T. **Furtenbacher**, P. **Árendás**, G. Mellau, and A. G. **Császár**, *Sci. Rep.* **2014**, *4*, 4654.

(f) A. G. **Császár**, T. **Furtenbacher**, and P. **Árendás**, *J. Phys. Chem. A* **2016**, *120*, 8949-8969.

(g) P. **Árendás**, T. **Furtenbacher**, and A. G. **Császár**, *J. Math. Chem.* **2016**, *54*, 806-822.

(h) **R. Tóbiás**, T. **Furtenbacher**, and A. G. **Császár**, *J. Quant. Spectrosc. Rad. Transfer* **2017**, *203*, 557-564.

Fundamental definitions

Definition 1: $N_S = \langle L, T, I, \varsigma \rangle$ is a **weighted, directed multigraph** if

(a) L is the set of vertices,

(b) T is the set of edges,

(c) $I : T \to L \times L$ is the **incidence function** of the edges, where $L \times L$ is the set of ordered pairs of L, and

(d) $\varsigma : T \to \mathbb{R}^+$ is a weight function over the set T.

Definition 2: The quadruple $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ in Def. 1 is a **spectroscopic network** (SN) if

(a) L is the set of (rovibronic) **energy levels**,

(b) T is the set of **transitions** among the energy levels,

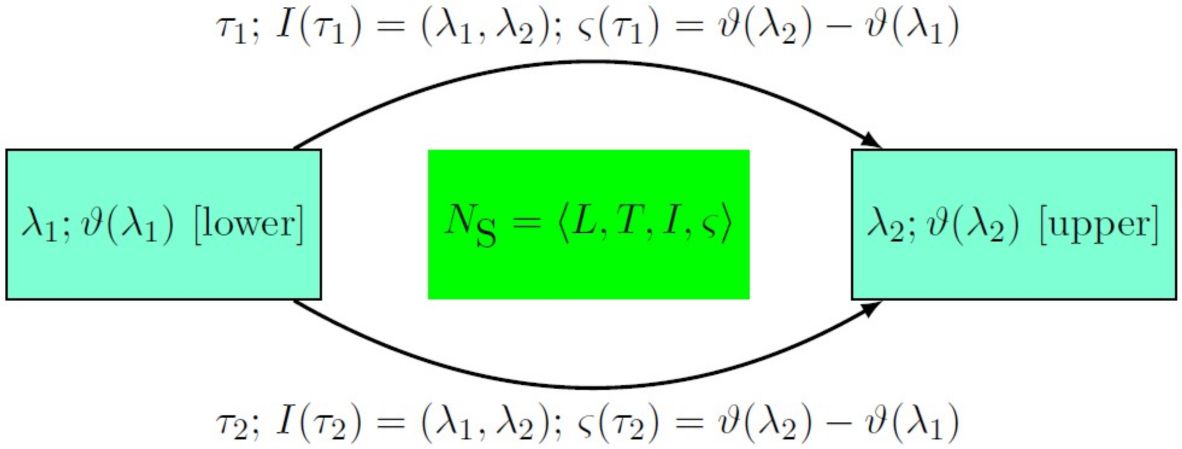
(c) there is a **(conservative) potential function** $\vartheta : L \to \mathbb{R}$ such that, for all $\tau \in T$ of incidence $I(\tau) = (\lambda_1, \lambda_2)$, $\varsigma(\tau) = \vartheta(\lambda_2) - \vartheta(\lambda_1)$ (**Rydberg–Ritz principle**), and

(d) ς is the **function** of **potential differences**.

Definition 3: If $\tau \in T$ with $I(\tau) = (\lambda_1, \lambda_2)$, then λ_1 and λ_2 are the **lower and upper levels** of τ, respectively. λ_1 and λ_2 will be denoted with low(τ) and up(τ), respectively, where low: $T \to L$ and up: $T \to L$ are two functions.

Definition 4: The transitions $\tau' \in T$ and $\tau'' \in T$ are **coincident** if $I(\tau') = I(\tau'')$.

Definition 5: The set of all $\tau' \in T$ coincident to $\tau \in T$ is the **coincidence class** of τ.



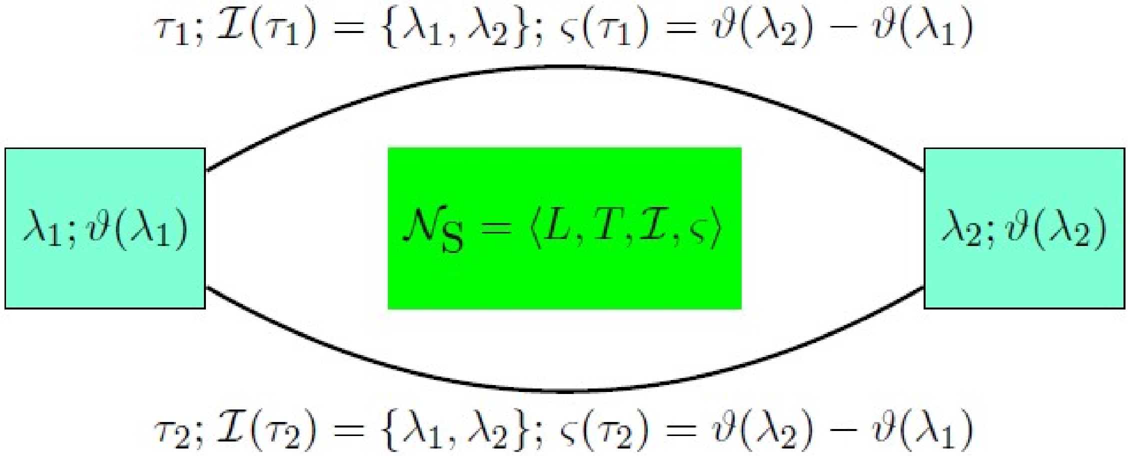
$$L = \{\lambda_1, \lambda_2\}$$
$$T = \{\tau_1, \tau_2\}$$
$$L \times L = \{(\lambda_1, \lambda_1), (\lambda_1, \lambda_2), (\lambda_2, \lambda_1), (\lambda_2, \lambda_2)\}$$
$$I : T \to L \times L$$
$$\varsigma : T \to \mathbb{R}^+$$
$$\vartheta : L \to \mathbb{R}$$

Figure 1: An example for a SN

Definition 6: $\mathcal{N}_{\mathrm{S}} = \langle L, T, \mathcal{I}, \varsigma \rangle$ is the **underlying network** of $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$, if

(a) $L \bullet L$ is the set of L's unordered pairs, and

(b) $\mathcal{I} : T \to L \bullet L$ is the **undirected incidence function** with $\mathcal{I}(\tau) = \{\lambda_1, \lambda_2\}$ for all $\tau \in T$ with $I(\tau) = (\lambda_1, \lambda_2)$.



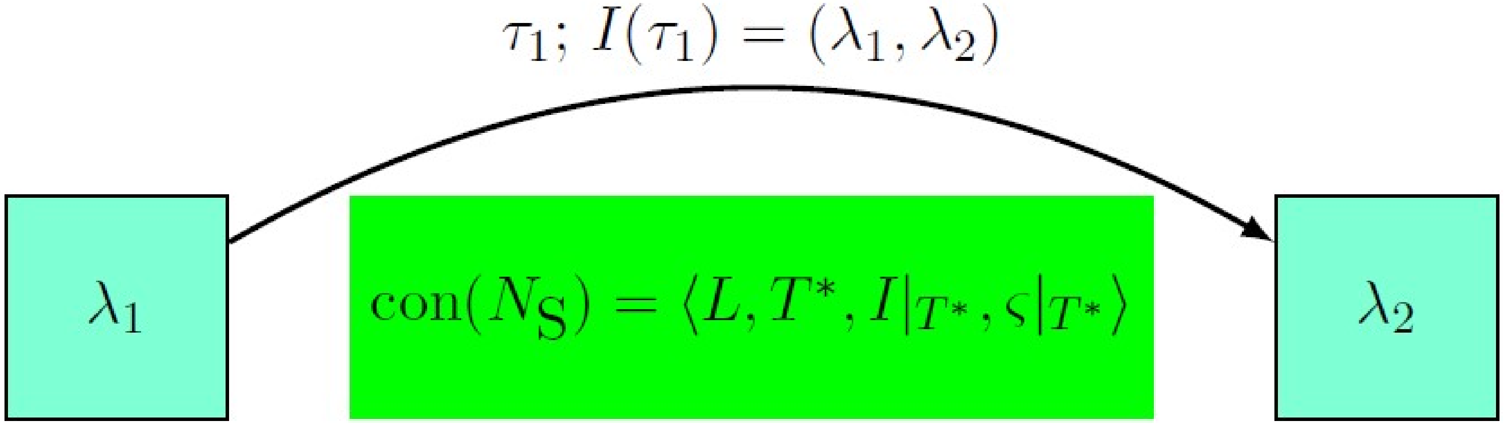
$$L = \{\lambda_1, \lambda_2\}$$
$$T = \{\tau_1, \tau_2\}$$
$$L \bullet L = \{\{\lambda_1\}, \{\lambda_2\}, \{\lambda_1, \lambda_2\}\}$$
$$\mathcal{I} : T \to L \bullet L$$
$$\varsigma : T \to \mathbb{R}^+$$
$$\vartheta : L \to \mathbb{R}$$

Figure 2: The underlying spectroscopic network of the SN of Fig. 1

Definition 7: $\mathrm{sub}(N_\mathrm{S}) = \left\langle L_\mathrm{sub}, T_\mathrm{sub}, I\big|_{T_\mathrm{sub}}, \varsigma\big|_{T_\mathrm{sub}} \right\rangle$ is a **subnet-work** of $N_\mathrm{S} = \langle L, T, I, \varsigma \rangle$ if $T_\mathrm{sub} \subseteq T$, $L_\mathrm{sub} = \cup_{\tau \in T_\mathrm{sub}} \mathcal{I}(\tau)$, and $I\big|_{T_\mathrm{sub}}$ and $\varsigma\big|_{T_\mathrm{sub}}$ are restrictions of I and ς to T_sub, respectively.

Definition 8: $T^* \subseteq T$ is a **set of unique transitions** if T^* is one of the largest subset of T for which $I\big|_{T^*}$ is injective.

Definition 9: $\mathrm{con}(N_\mathrm{S}) = \left\langle L, T^*, I\big|_{T^*}, \varsigma\big|_{T^*} \right\rangle$ is a **contraction** of $N_\mathrm{S} = \langle L, T, I, \varsigma \rangle$ if T^* is a set of unique transitions.



$$L = \{\lambda_1, \lambda_2\}$$
$$T^* = \{\tau_1\}$$
$$L \times L = \{(\lambda_1, \lambda_1), (\lambda_1, \lambda_2), (\lambda_2, \lambda_1), (\lambda_2, \lambda_2)\}$$
$$I : T \to L \times L$$

Figure 3: The contraction graph of the SN of Fig. 1

Definition 10: $\{\tau_1, \tau_2, \ldots, \tau_L\} \subseteq T$ is an **oriented path** of length $\mathcal{L}$ from λ_1 to $\lambda_{\mathcal{L}+1}$ in $N_S = \langle L, T, I, \varsigma \rangle$ if

$$I(\tau_1) = (\lambda_1, \lambda_2)$$

$$I(\tau_2) = (\lambda_2, \lambda_3)$$

$$\vdots$$

$$I(\tau_{\mathcal{L}}) = (\lambda_{\mathcal{L}}, \lambda_{\mathcal{L}+1}).$$

Definition 11: $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}}\} \subseteq T$ is a **path** of length $\mathcal{L}$ between λ_1 and $\lambda_{\mathcal{L}+1}$ in $N_S = \langle L, T, I, \varsigma \rangle$ if

$$\mathcal{I}(\tau_1) = \{\lambda_1, \lambda_2\}$$

$$\mathcal{I}(\tau_2) = \{\lambda_2, \lambda_3\}$$

$$\vdots$$

$$\mathcal{I}(\tau_{\mathcal{L}}) = \{\lambda_{\mathcal{L}}, \lambda_{\mathcal{L}+1}\}.$$

Definition 12: $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ is **connected** if there is a path between every pair of energy levels.

Definition 13: $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}}\} \subseteq T$ is an **oriented cycle** of length $\mathcal{L}$ in $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ if $I(\tau_{\mathcal{L}}) = (\lambda_{\mathcal{L}}, \lambda_1)$ and $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}-1}\}$ is an oriented path from λ_1 to $\lambda_{\mathcal{L}}$.

Definition 14: $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}}\} \subseteq T$ is a **cycle** of length $\mathcal{L}$ in $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ if $\mathcal{I}(\tau_{\mathcal{L}}) = \{\lambda_1, \lambda_{\mathcal{L}}\}$ and $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}-1}\}$ is a path between λ_1 and $\lambda_{\mathcal{L}}$.

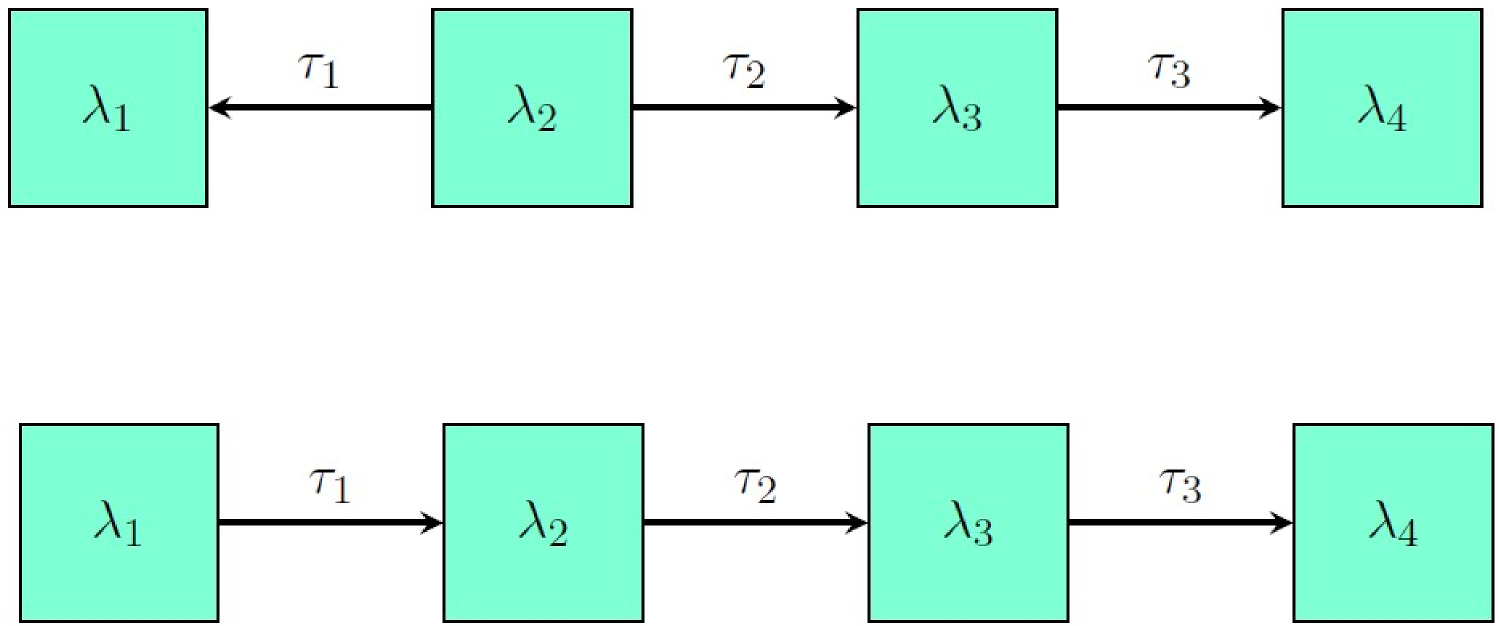


Figure 4: Examples for (a) a path and (b) an oriented path

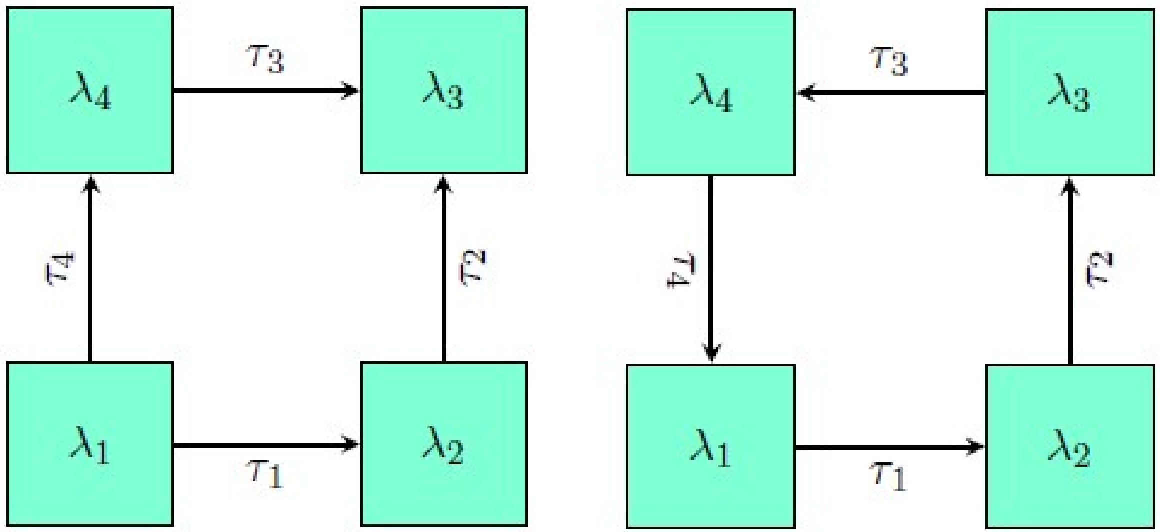


Figure 5: Examples for (a) a cycle and (b) an oriented cycle

Simple propositions related to SNs

Proposition 1: If $\tau \in T$ with $I(\tau) = (\lambda_1, \lambda_2)$, then there is no $\tau' \in T$ such that $I(\tau') = (\lambda_2, \lambda_1)$.

Proof:

Since $\varsigma(\tau) = \vartheta(\lambda_2) - \vartheta(\lambda_1) > 0$ for all $\tau \in T$, then $\varsigma(\tau') = \vartheta(\lambda_1) - \vartheta(\lambda_2) = -\varsigma(\tau) < 0$, which is impossible.

Proposition 2: If $\tau \in T$ with $I(\tau) = (\lambda_1, \lambda_2)$, then $\lambda_1 \neq \lambda_2$.

Proof:

If $\lambda_1 = \lambda_2$, then $\varsigma(\tau) = \vartheta(\lambda_1) - \vartheta(\lambda_1) = 0$, which is a contradiction because of $\varsigma(\tau) > 0$.

Proposition 3: If $\tau' \in T$ and $\tau'' \in T$ are coincident, *i.e.*, $I(\tau') = I(\tau'') = (\lambda_1, \lambda_2)$, then

$$\varsigma(\tau') = \varsigma(\tau'') = \vartheta(\lambda_2) - \vartheta(\lambda_1).$$

Proposition 4: If ϑ and ϑ' are two conservative potential functions of ς, such that $\vartheta'(\lambda)=\vartheta(\lambda)+c$ with a $c\in\mathbb{R}$ for all $\lambda\in L$, then

$$\varsigma(\tau)=\vartheta(\lambda_2)-\vartheta(\lambda_1)=\left(\vartheta(\lambda_2)+c\right)-\left(\vartheta(\lambda_1)+c\right)$$

$$=\vartheta'(\lambda_2)-\vartheta'(\lambda_1)$$

Since $c=-\min_{\lambda\in L}\vartheta(\lambda)$ implies $\min_{\lambda\in L}\vartheta'(\lambda)=0$, we can suppose that $\min_{\lambda\in L}\vartheta(\lambda)=0$.

Proposition 5: If $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}}\} \subseteq T$ is an oriented path with $I(\tau_i) = (\lambda_i, \lambda_{i+1})$ $(1 \le i \le \mathcal{L})$, then

$$\varsigma(\tau_1) + \varsigma(\tau_2) + \ldots + \varsigma(\tau_{\mathcal{L}}) =$$

$$\vartheta(\lambda_2) - \vartheta(\lambda_1) + \vartheta(\lambda_3) - \vartheta(\lambda_2) + \ldots + \vartheta(\lambda_{\mathcal{L}+1}) - \vartheta(\lambda_{\mathcal{L}}) =$$

$$\vartheta(\lambda_{\mathcal{L}+1}) - \vartheta(\lambda_1).$$

Proposition 6: If $\Theta = \{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}}\} \subseteq T$ is a path with $\mathcal{I}(\tau_i) = \{\lambda_i, \lambda_{i+1}\}$ $(1 \le i \le \mathcal{L})$, then there is a $\Phi_\Theta : \Theta \to \{-1, 1\}$ **spectral sign function** with $\phi_i = \Phi_\Theta(\tau_i)$ such that

$$\phi_1 \varsigma(\tau_1) + \phi_2 \varsigma(\tau_2) + \ldots + \phi_{\mathcal{L}} \varsigma(\tau_{\mathcal{L}}) = \vartheta(\lambda_{\mathcal{L}+1}) - \vartheta(\lambda_1).$$

Proposition 7: If $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}}\} \subseteq T$ is an oriented cycle with $I(\tau_i) = (\lambda_i, \lambda_{i+1})$ $(1 \le i \le \mathcal{L} - 1)$ and $I(\tau_{\mathcal{L}}) = (\lambda_{\mathcal{L}}, \lambda_1)$, then

$$\varsigma(\tau_1) + \varsigma(\tau_2) + \ldots + \varsigma(\tau_{\mathcal{L}}) = 0.$$

Proof:

Since $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}-1}\} \subseteq T$ is an oriented path, then

$$\varsigma(\tau_1) + \varsigma(\tau_2) + \ldots + \varsigma(\tau_{\mathcal{L}-1}) = \vartheta(\lambda_{\mathcal{L}}) - \vartheta(\lambda_1) = -\varsigma(\tau_{\mathcal{L}}),$$

utilizing $\varsigma(\tau_{\mathcal{L}}) = \vartheta(\lambda_1) - \vartheta(\lambda_{\mathcal{L}})$, that is,

$$\varsigma(\tau_1) + \varsigma(\tau_2) + \ldots + \varsigma(\tau_{\mathcal{L}-1}) + \varsigma(\tau_{\mathcal{L}}) = 0.$$

Proposition 8: If $\varTheta = \{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}}\} \subseteq T$ is a cycle with $\mathcal{I}(\tau_i)$ $(1 \le i \le \mathcal{L} - 1)$ and $\mathcal{I}(\tau_{\mathcal{L}}) = \{\lambda_{\mathcal{L}}, \lambda_1\}$, then there is a $\varPhi_{\varTheta} : \varTheta \to \{-1, 1\}$ function with $\phi_i = \varPhi_{\varTheta}(\tau_i)$ such that

$$\phi_1 \varsigma(\tau_1) + \phi_2 \varsigma(\tau_2) + \ldots + \phi_{\mathcal{L}} \varsigma(\tau_{\mathcal{L}}) = 0 \; (*).$$

Proposition 9: If $\phi_1, \phi_2, \ldots, \phi_{\mathcal{L}} \in \{1, -1\}$ satisfy (*), then $-\phi_1, -\phi_2, \ldots, -\phi_{\mathcal{L}}$ also satisfy (*).

Proposition 10: A cycle $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}}\} \subseteq T$ is oriented iff $\phi_1 = \phi_2 = \ldots = \phi_{\mathcal{L}}$ obeying (*).

Proposition 11: If $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}}\} \subseteq T$ is a cycle with $\mathcal{I}(\tau_i) = \{\lambda_i, \lambda_{i+1}\}$ ($1 \le i \le \mathcal{L} - 1$) and $\mathcal{I}(\tau_{\mathcal{L}}) = \{\lambda_{\mathcal{L}}, \lambda_1\}$, then it is not oriented.

Proof:

Suppose that $\phi = \phi_1 = \phi_2 = \ldots = \phi_{\mathcal{L}} \in \{1, -1\}$ meets (*), *i.e.*,

$$\phi\left[\varsigma(\tau_1) + \varsigma(\tau_2) + \ldots + \varsigma(\tau_{\mathcal{L}-1}) + \varsigma(\tau_{\mathcal{L}})\right] = 0.$$

Then, we get a contradiction due to $\varsigma(\tau_i) > 0$ ($1 \le i \le \mathcal{L}$):

$$0 < \varsigma(\tau_1) + \varsigma(\tau_2) + \ldots + \varsigma(\tau_{\mathcal{L}-1}) = -\varsigma(\tau_{\mathcal{L}}) < 0.$$

Proposition 12: If $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ is a connected graph and $\vartheta(\lambda_0) = 0$ with a $\lambda_0 \in L$, then the $\vartheta(\lambda)$ values are uniquely determined by $\varsigma : T \to \mathbb{R}$ for all $\lambda \in L$.

Proof:

Since there is a $\{\tau_1, \tau_2, \ldots, \tau_{\mathcal{L}}\} \subseteq T$ path between λ_0 and an arbitrary $\lambda \in L$ with $\lambda \neq \lambda_0$, the following relation holds:

$$\vartheta(\lambda) - \vartheta(\lambda_0) = \vartheta(\lambda) = \phi_1 \varsigma(\tau_1) + \phi_2 \varsigma(\tau_2) + \ldots + \phi_{\mathcal{L}} \varsigma(\tau_{\mathcal{L}})$$

with proper $\phi_1, \phi_2, \ldots, \phi_{\mathcal{L}} \in \{1, -1\}$ signs.

Experimental realizations of SNs

Definition 15: $R_{\mathrm{S}} = \langle L, T, I, \sigma, \delta \rangle$ is a(n experimental) **realization** of the network $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ if $\sigma : T \to \mathbb{R}$, $\varepsilon : T \to \mathbb{R}$, and $\delta : T \to \mathbb{R}$ are functions, furthermore,

(a) $\varepsilon(\tau)$ is a random variable with zero expectation value and $\delta(\tau)$ standard deviation, and

(b) $\sigma(\tau) = \varsigma(\tau) + \varepsilon(\tau)$

for all $\tau \in T$.

Definition 16: $\mathrm{con}\left(R_{\mathrm{S}}\right)=\left\langle L,T^{*},I\big|_{T^{*}},\sigma\big|_{T^{*}},\delta\big|_{T^{*}}\right\rangle$ is a **contraction** of the realization $R_{\mathrm{S}}=\left\langle L,T,I,\sigma,\delta\right\rangle$ if T^{*} is a set of unique transitions.

Definition 17: Let $E_{\vartheta(\lambda)}$ be an **estimator** based on the realization $R_{\mathrm{S}}=\left\langle L,T,I,\sigma,\delta\right\rangle$ for a $\lambda\in L$, and $\vartheta_{\mathrm{opt}}\left(\lambda\right)$ is the **estimation** for $\vartheta\left(\lambda\right)$ using $E_{\vartheta(\lambda)}$. The function $\vartheta_{\mathrm{opt}}:L\rightarrow\mathbb{R}^{+}$ is an **optimal estimation** of ϑ if $E_{\vartheta(\lambda)}$ is the **best unbiased**

estimator of $\vartheta(\lambda)$ for all $\lambda \in L$, *i.e.*, the expectation value of $E_{\vartheta(\lambda)}$ is $\vartheta(\lambda)$ and the variance of $E_{\vartheta(\lambda)}$ is minimal.

Proposition 13: For a given $R_{\mathrm{S}} = \langle L, T, I, \sigma, \delta \rangle$ realization

$$\vartheta_{\mathrm{opt}} = \arg\min_{\vartheta} S(\vartheta),$$

where $S(\vartheta)$ is the following objective function:

$$S(\vartheta) = \sum_{\tau \in T} \frac{1}{\delta^2(\tau)} \Big(\sigma(\tau) - \vartheta\big(\mathrm{up}(\tau)\big) + \vartheta\big(\mathrm{low}(\tau)\big)\Big)^2 .$$

(MARVEL procedure; see also (a) and (b) in Sec. „Related studies")

Inconsistencies in $R_S = \langle L, T, I, \sigma, \delta \rangle$

(a) According to the „source" of the problem:

– (transcription) errors

– (measurement) inaccuracies (rel. unc. is less than $\mathbf{10^{-5}}$)

(b) Based on the „location" of the problem ($\tau \in T$):

– error or inaccuracy among the $I(\tau)$ values

– error or inaccuracy among the $\sigma(\tau)$ values

– error or inaccuracy among the $\delta(\tau)$ values

Techniques to treat inconsistencies

(a) observe the trends of the residuals

$$\rho(\tau) = \sigma(\tau) - \vartheta_{\text{opt}}\big(\text{up}(\tau)\big) + \vartheta_{\text{opt}}\big(\text{low}(\tau)\big)$$

for all $\tau \in T$,

(b) adjust the values of δ in a "reweighting procedure",

(c) use cycle bases to identify incorrect cycles (ECART),

(c) and seek for further spectroscopic information on the possibly incorrect transitions.

Cycle bases

Definition 18: $\mathcal{C}$ is the **cycle space** of $N_\mathrm{S} = \langle L, T, I, \varsigma \rangle$ if it contains all the cycles $C \subseteq T$ in $N_\mathrm{S} = \langle L, T, I, \varsigma \rangle$.

Definition 19: $C' \in \mathcal{C}$ and $C'' \in \mathcal{C}$ are **independent** if their $C' \Delta C''$ symmetric difference is a non-empty set, that is,

$$C' \Delta C'' = \left(C' \setminus C''\right) \cup \left(C'' \setminus C'\right) \neq \varnothing,$$

otherwise C' and C'' are called **dependent**.

Definition 20: $C', C'' \in \mathcal{C}$ are **disjoint** if $C' \cap C'' = \varnothing$.

Proposition 14: If the non-disjoint cycles $C', C'' \in \mathcal{C}$ are independent, then $C' \Delta C'' \in \mathcal{C}$.

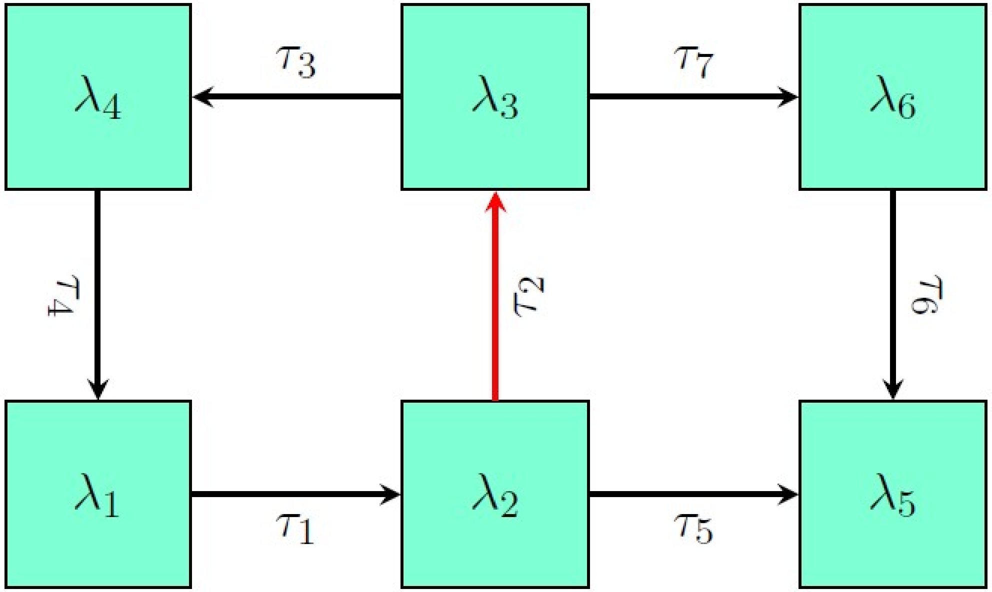


Figure 6: Symmetric difference of two cycles of length 4 obtained by leaving the transition τ_2

Definition 21: $\mathcal{C}_{\mathrm{B}} \subseteq \mathcal{C}$ is a **cycle basis** of the cycle space $\mathcal{C}$ if there is a set $\mathcal{X} \subseteq \mathcal{C}_{\mathrm{B}}$ with independent cycles such that either $C \in \mathcal{X}$ or $C = \Delta_{\chi \in \mathcal{X}} X$ for each $C \in \mathcal{C}$. The entries of $\mathcal{C}_{\mathrm{B}}$ are called **basic cycles**.

Proposition 15: Each cycle space has a cycle basis.

Definition 22: $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ is **acyclic** if $\mathcal{C} = \varnothing$.

Definition 23: $N_{\mathrm{S}}^{[\![\tilde{T}]\!]} = \langle L, \tilde{T}, I|_{\tilde{T}}, \varsigma|_{\tilde{T}} \rangle$ is a **spanning tree** of $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ if $N_{\mathrm{S}}^{[\![\tilde{T}]\!]}$ is (a) acyclic, (b) connected, and (c) $\cup_{\tau \in \tilde{T}} \mathcal{I}(\tau) = L$.

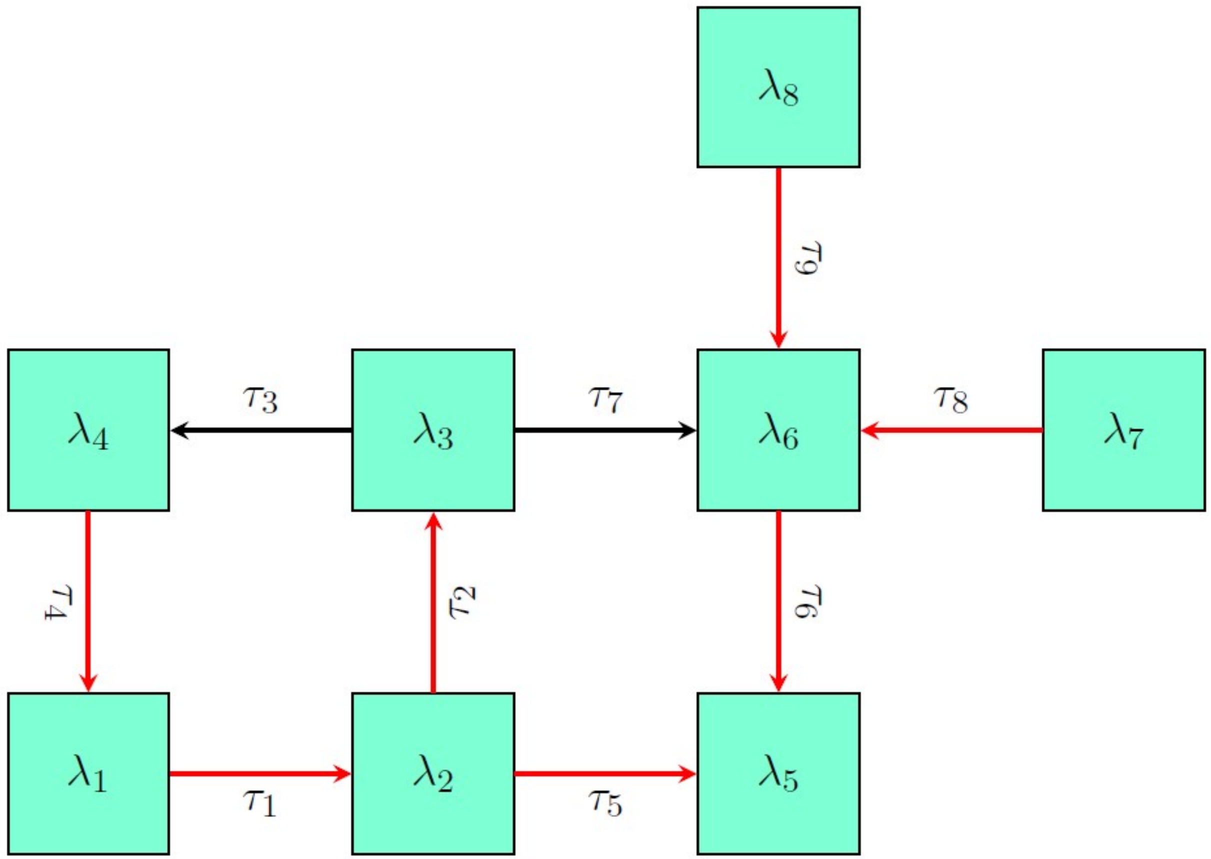


Figure 7: A spanning tree (red edges) of a SN with 7 energy levels

Proposition 16: If $N_{\mathrm{S}}^{[\![\tilde{T}]\!]} = \left\langle L, \tilde{T}, I\big|_{\tilde{T}}, \varsigma\big|_{\tilde{T}} \right\rangle$ is a spanning tree of $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ and $\tau \in T \setminus \tilde{T}$ with $\mathcal{I}(\tau) = \{\lambda_1, \lambda_2\}$, then $C_\tau = P_\tau \cup \{\tau\}$ a so-called **fundamental cycle** of N_{S}, where $P_\tau \subseteq \tilde{T}$ is a unique path in $N_{\mathrm{S}}^{[\![\tilde{T}]\!]}$ between λ_1 and λ_2.

Proposition 17: If $\tau' \neq \tau''$, then $C_{\tau'}, C_{\tau''} \in \mathcal{C}$ are independent.

Proposition 18: If $\operatorname{con}(N_{\mathrm{S}}) = N_{\mathrm{S}}$, and N_{S} is a connected, then $\mathcal{C}_{\tilde{T}} = \left\{ C_\tau \in \mathcal{C} : \tau \in T \setminus \tilde{T} \right\}$ is a cycle basis of $\mathcal{C}$ with $\left|\mathcal{C}_{\tilde{T}}\right| = |T| - |L| + 1$.

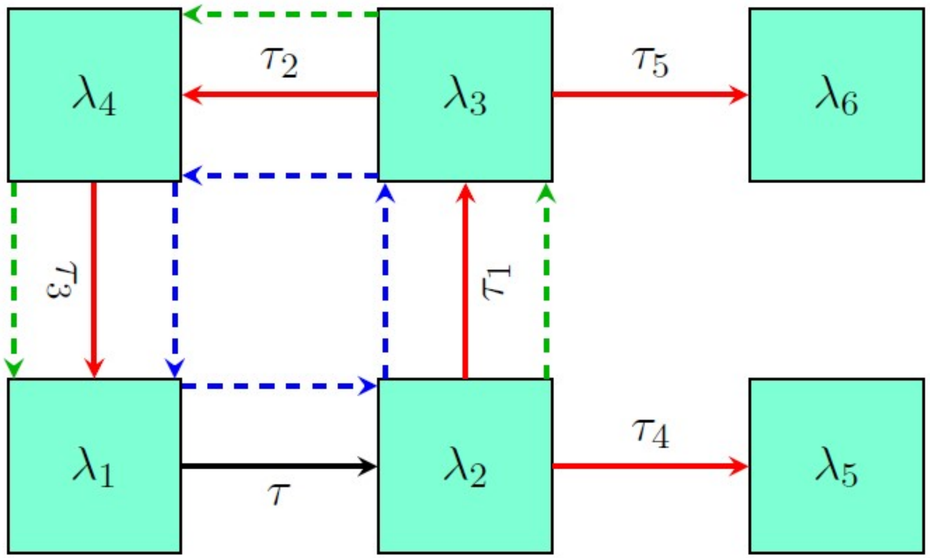


Figure 8: A fundamental cycle associated with τ

(**red:** spanning tree edges; **black:** non-spanning tree edges; **green:** auxiliary edges denoting the path between λ_1 and λ_2 in the spanning tree; **blue:** auxiliary edges denoting the fundamental cycle associated to τ)

ECART algorithm
(see also (d) in Sec. „Related studies”)

(a) construct a cycle basis ($\mathcal{C}_{\mathrm{B}}$),

(b) calculate the **discrepancy** of each basic cycle $\chi \in \mathcal{C}_{\mathrm{B}}$, $\mathcal{D}(\chi)$, using the spectral sign function $\Phi_\chi : \chi \to \{-1,1\}$ as

$$\mathcal{D}(\chi) = \left|\sum_{\tau\in\chi} \Phi_\chi(\tau)\sigma(\tau)\right|$$

(c) determine the $\mathcal{T}(\chi)$ **threshold** of each $\chi \in \mathcal{C}_{\mathrm{B}}$ as

$$\mathcal{T}(\chi) = \sum_{\tau\in\chi} \delta(\tau),$$

(d) denote the basic cycles $\chi \in \mathcal{C}_{\mathrm{B}}$ as **bad** if

$$\mathcal{D}(\chi) - \mathcal{T}(\chi) \geq p_{\text{cut-off}}(\chi),$$

where $p_{\text{cut-off}}(\chi) \in \mathbb{R}^{+}$ is a cut-off parameter associated to the cycle χ,

(e) denote the transitions $\tau \in \cup_{\chi \in \mathcal{C}_{\mathrm{B}}} \chi$ as **suspicious** or **harmless** depending on whether it is included in at least one bad basic cycle or not, respectively, and

(f) check the suspicious transitions.

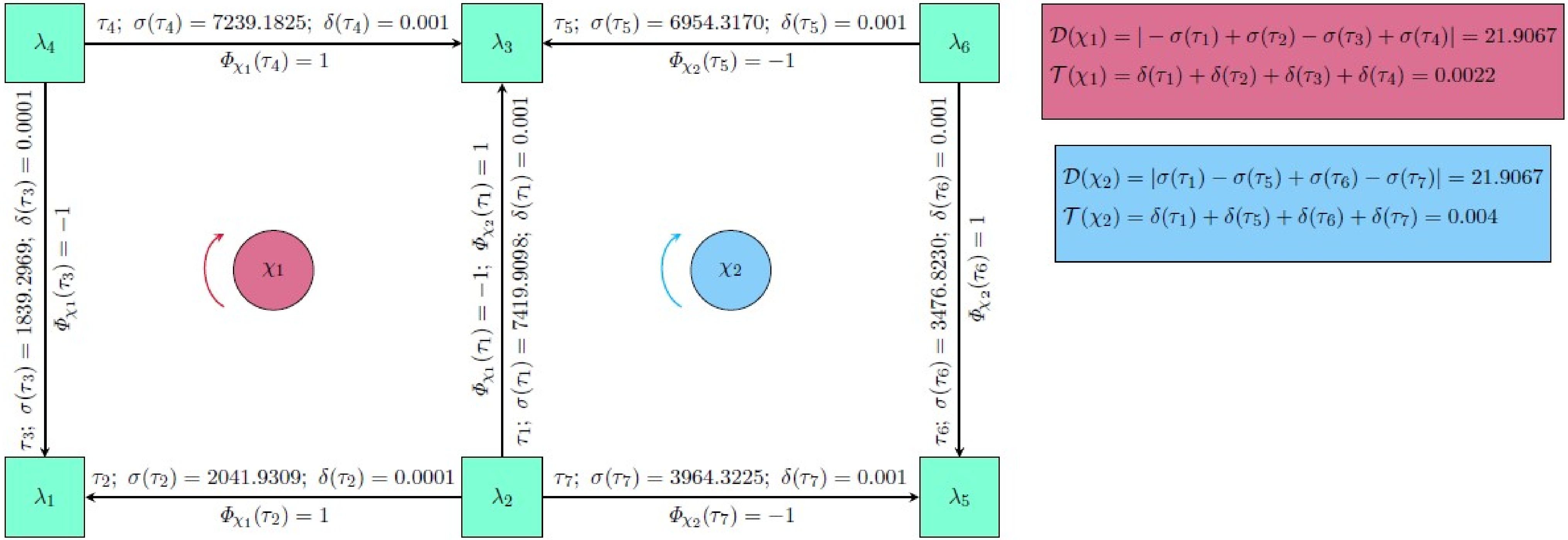


Figure 9: Two bad basic cycles, χ_1 and χ_2. The spectral signs are determined based on the red and blue directions. The large and nearly identical discrepancies are due to the $\sigma(\tau_1)$ value common in χ_1 and χ_2.

Minimum Cycle Basis (MCB)

Definition 24: The **total cost** of a $\mathcal{C}_\mathrm{B} \subseteq \mathcal{C}$, denoted with $\kappa_\mathrm{tot}\left(\mathcal{C}_\mathrm{B}\right)$ is $\sum_{\chi \in \mathcal{C}_\mathrm{B}} |\chi|$, where $|\chi|$ is the length of the cycle χ.

Definition 25: $\mathcal{C}_\mathrm{B,min}$ is a **minimum cycle basis** of $\mathcal{C}$ if

$$\mathcal{C}_\mathrm{B,min} = \arg\min_{\mathcal{C}_\mathrm{B}} \kappa_\mathrm{tot}\left(\mathcal{C}_\mathrm{B}\right).$$

Proposition 19: If $\mathrm{con}\left(N_\mathrm{S}\right) = N_\mathrm{S}$, and N_S is connected, $\mathcal{C}_\mathrm{B,min}$ can be built with $O\left(|T|^2 |L| + |T||L|^2 \log\left(|L|\right)\right)$ complexity.

Advantages and disadvantages of MCBs

(a) extremely transparent due to their short cycles,

(b) short cycles are more sensitive to the errors and the inaccuracies than longer cycles,

(c) MCBs require larger computational expense than the cycle bases determined with the well-known spanning-tree-search techniques.

Open problems

Definition 26: Let $\left\langle L, T^{*}, I\big|_{T^{*}}, \varsigma\big|_{T^{*}} \right\rangle$ and $\left\langle L, T'^{*}, I\big|_{T'^{*}}, \varsigma\big|_{T'^{*}} \right\rangle$ be two contractions of $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ with their cycle bases $\mathcal{C}_{\mathrm{B}}^{*}$ and $\mathcal{C}_{\mathrm{B}}'^{*}$, respectively. $\mathcal{C}_{\mathrm{B}}^{*}$ and $\mathcal{C}_{\mathrm{B}}'^{*}$ are **congruent** ($\mathcal{C}_{\mathrm{B}}^{*} \simeq \mathcal{C}_{\mathrm{B}}'^{*}$) if there is a bijection $f : \mathcal{C}_{\mathrm{B}}^{*} \leftrightarrow \mathcal{C}_{\mathrm{B}}'^{*}$ such that for each $\chi \in \mathcal{C}_{\mathrm{B}}'^{*}$ we have $I\big|_{\chi} = I\big|_{f(\chi)}$.

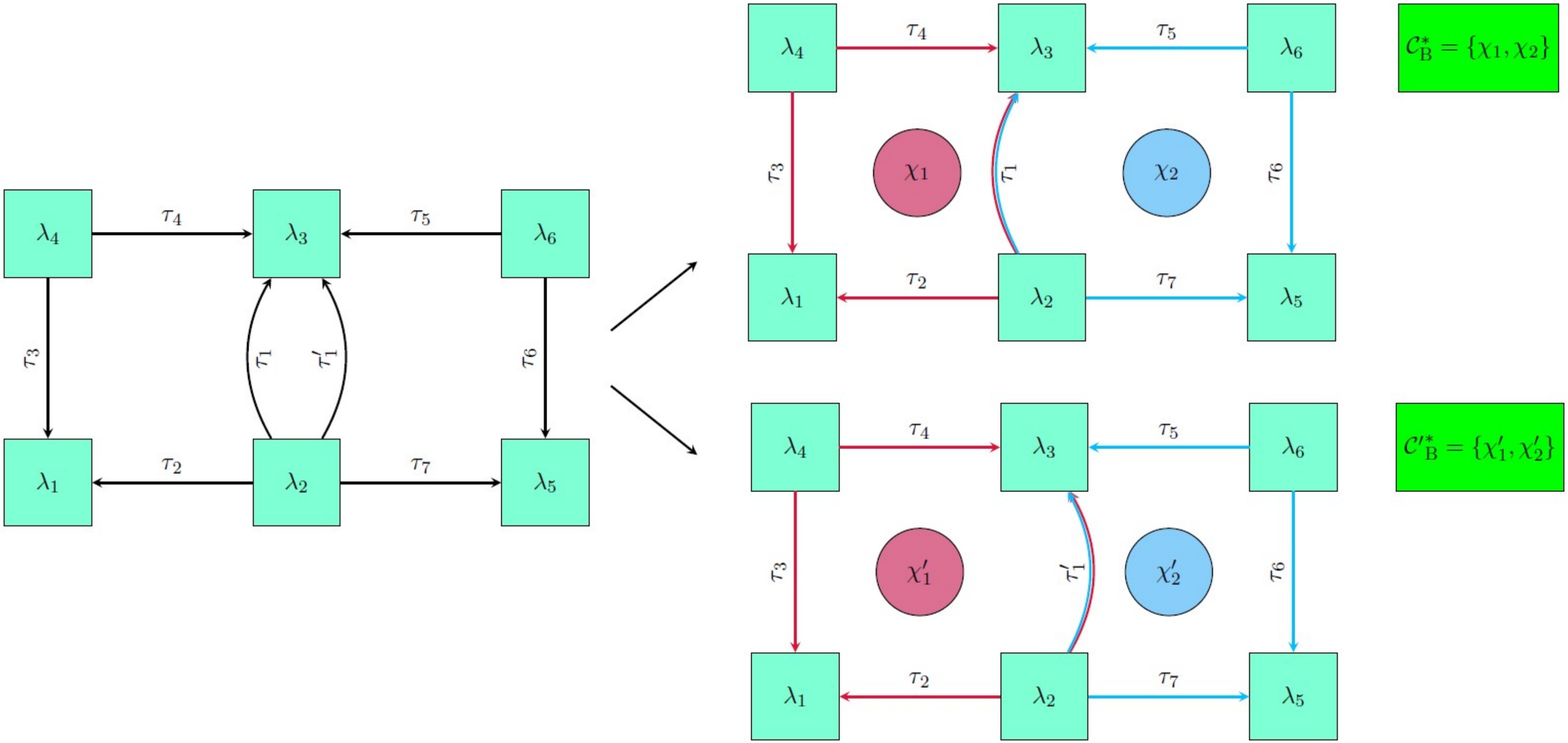


Figure 10: Congruent cycle bases ($\mathcal{C}^*_{\mathrm{B}} \simeq \mathcal{C}'^*_{\mathrm{B}}$)

Problem 1: Let $\tilde{N}_{\mathrm{S}}^{*}=\left\langle L,\tilde{T}^{*},I\big|_{\tilde{T}^{*}},\varsigma\big|_{\tilde{T}^{*}}\right\rangle$ be a contraction of $N_{\mathrm{S}}=\langle L,T,I,\varsigma\rangle$ with the cycle basis $\tilde{\mathcal{C}}_{\mathrm{B}}^{*}$. Let us ask how to find a $N_{\mathrm{S,opt}}^{*}=\left\langle L,T_{\mathrm{opt}}^{*},I\big|_{T_{\mathrm{opt}}^{*}},\varsigma\big|_{T_{\mathrm{opt}}^{*}}\right\rangle$ **optimal contraction** with

$$N_{\mathrm{S,opt}}^{*}=\underset{\substack{N_{\mathrm{S}}^{*}\\ \mathcal{C}_{\mathrm{B}}^{*}\simeq\tilde{\mathcal{C}}_{\mathrm{B}}^{*}}}{\arg\min}\left(\sum_{\chi\in\mathcal{C}_{\mathrm{B}}^{*}}\mathcal{D}(\chi)\right),$$

where $N_{\mathrm{S}}^{*}=\left\langle L,T^{*},I\big|_{T^{*}},\varsigma\big|_{T^{*}}\right\rangle$ is a contraction of N_{S} with the cycle basis, $\mathcal{C}_{\mathrm{B}}^{*}$, congruent to $\tilde{\mathcal{C}}_{\mathrm{B}}^{*}$, and $\mathcal{D}(\chi)$ is the discrepancy of $\chi\in\mathcal{C}_{\mathrm{B}}^{*}$.

Problem 2: Let $N_{\mathrm{S}} = \langle L, T, I, \varsigma \rangle$ be a connected network. Using a spanning tree/cycle basis of N_{S}, the question is how to decide wheather the network $\left\langle L, T \setminus \tau, I|_{T\setminus\tau}, \varsigma|_{T\setminus\tau} \right\rangle$ is **disconnected** for a $\tau \in T$.

Problem 3: Considering a $\left\langle L, T^{*}, I|_{T^{*}}, \sigma|_{T^{*}}, \delta|_{T^{*}} \right\rangle$ contraction of $R_{\mathrm{S}} = \langle L, T, I, \sigma, \delta \rangle$, it is unclear how to determine one of the largest **"clear" subset** of T^{*}, denoted with T_{c}^{*}, for which $\sigma|_{T_{\mathrm{c}}^{*}}$ does not contain outliers.

Problem 4: Let $R_{\mathrm{S}}=\langle L,T,I,\sigma,\delta\rangle$ be a realization of $N_{\mathrm{S}}=\langle L,T,I,\varsigma\rangle$. How to give a good estimation for the standard deviations $\delta(\tau)$ ($\tau\in T$) using a cycle basis of N_{S}?

Problem 5: Let $N_{\mathrm{S}}=\langle L,T,I,\varsigma\rangle$ be a network with $\mathrm{con}(N_{\mathrm{S}})=N_{\mathrm{S}}$. How can one obtain a cycle basis $\mathcal{C}_{\mathrm{B,max}}$ of N_{S} such that

$$\mathcal{C}_{\mathrm{B,max}}=\arg\max_{\mathcal{C}_{\mathrm{B}}}\sum_{\chi\in\mathcal{C}_{\mathrm{B}}}\mathcal{D}(\chi).$$

where $\mathcal{D}(\chi)$ is the discrepancy of χ.

Thank you for your attention!