Cluster Expansions and Connected Graph Identity

Mei Yin

Department of Mathematics
University of Arizona

April 9, 2008

First, I will give a relatively gentle introduction to cluster expansions, which refers to a technique for controlling measures on spaces of large dimension that are defined by a density with respect to a reference measure. After that, I will present a connected graph identity that may lead to an alternative proof of the Fernández-Procacci result.

The following terminologies are common in the area of mathematical physics.

- The setting for the problem is a finite set $\mathcal{P}$. Each element of $\mathcal{P}$ is a *location* at which some number of particles may be present.
- For each pair p, q of locations we have an interaction potential energy $V(p, q)$. We suppose that $V(p, q) = V(q, p)$. Furthermore, we require that $0 \leq V(p, q) \leq +\infty$.
- In some applications each location may have additional internal structure. In such a context a location is often referred to as a *polymer*. And the energy $V(p, q)$ is assumed only the two values 0 and $+\infty$.

The following terminologies are common in the area of mathematical physics.

- The setting for the problem is a finite set $\mathcal{P}$. Each element of $\mathcal{P}$ is a *location* at which some number of particles may be present.
- For each pair p, q of locations we have an interaction potential energy $V(p, q)$. We suppose that $V(p, q) = V(q, p)$. Furthermore, we require that $0 \leq V(p, q) \leq +\infty$.
- In some applications each location may have additional internal structure. In such a context a location is often referred to as a *polymer*. And the energy $V(p, q)$ is assumed only the two values 0 and $+\infty$.

The following terminologies are common in the area of mathematical physics.

- The setting for the problem is a finite set $\mathcal{P}$. Each element of $\mathcal{P}$ is a *location* at which some number of particles may be present.
- For each pair p, q of locations we have an interaction potential energy $V(p, q)$. We suppose that $V(p, q) = V(q, p)$. Furthermore, we require that $0 \leq V(p, q) \leq +\infty$.
- In some applications each location may have additional internal structure. In such a context a location is often referred to as a *polymer*. And the energy $V(p, q)$ is assumed only the two values 0 and $+\infty$.

- The *Boltzmann factor* corresponding to the interaction potential energy is $\exp(-\beta V(p,q))$. Here the parameter $\beta > 0$ is actually the inverse temperature, measured in inverse energy units.
- The *interaction factor* $t(p,q) = \exp(-\beta V(p,q)) - 1$. Notice that $-1 \leq t(p,q) \leq 0$. The interaction factor will be the fundamental quantity later on.

- The *Boltzmann factor* corresponding to the interaction potential energy is $\exp(-\beta V(p,q))$. Here the parameter $\beta > 0$ is actually the inverse temperature, measured in inverse energy units.
- The *interaction factor* $t(p,q) = \exp(-\beta V(p,q)) - 1$. Notice that $-1 \leq t(p,q) \leq 0$. The interaction factor will be the fundamental quantity later on.

- One more ingredient is needed for the model. This is a set of parameters w_p , the *activity* parameters, associated to the locations p in $\mathcal{P}$. In the physical literature $w_p \geq 0$ is interpreted as a prior weight for the probability of finding a particle at location p.
- The variable of interest for us will be the *particle occupation number* N. This is a function from $\mathcal{P}$ to the natural numbers. Thus $N(p)$ is the number of particles at location p.

- One more ingredient is needed for the model. This is a set of parameters w_p , the *activity* parameters, associated to the locations p in $\mathcal{P}$. In the physical literature $w_p \geq 0$ is interpreted as a prior weight for the probability of finding a particle at location p.
- The variable of interest for us will be the *particle occupation number* N. This is a function from $\mathcal{P}$ to the natural numbers. Thus $N(p)$ is the number of particles at location p.

For each set $\{p, q\}$ of one or two locations, there is a corresponding number of pairs of particles given by $L_{\{p,q\}}(N) = N(p)N(q)$ if $p \neq q$ and by $L_{\{p,q\}}(N) = N(p)(N(p) - 1)/2$.

The *interaction product* associated with a particle occupation number is

$$Q(N) = \prod_{\{p,q\}} (1 + t(p, q))^{L_{\{p,q\}}(N)}.$$

This is just the product of the Boltzmann factors for all the pairs of particles.

For each set $\{p, q\}$ of one or two locations, there is a corresponding number of pairs of particles given by $L_{\{p,q\}}(N) = N(p)N(q)$ if $p \neq q$ and by $L_{\{p,q\}}(N) = N(p)(N(p) - 1)/2$.
The *interaction product* associated with a particle occupation number is

$$Q(N) = \prod_{\{p,q\}} (1 + t(p, q))^{L_{\{p,q\}}(N)}.$$

This is just the product of the Boltzmann factors for all the pairs of particles.

Write w^N for the product $\prod_q w_q^{N(q)}$. The probability model for an equilibrium gas is then

$$\mathrm{prob}(N) = \frac{1}{\Xi(w)} \frac{1}{N!} Q(N) w^N,$$

where $\Xi(w) = \sum_N \frac{1}{N!} Q(N) w^N$ is the normalizing factor, called the *grand partition function* in physics.
Notice that above, $N! = \prod_p N(p)!$ is a well-defined quantity.

Each probability of an individual particle occupation number function is tiny. Furthermore, it gets even smaller as the number of locations is increased. A quantity of physical interest that might be more robust is the expected number of particles at location p. This is given by

$$f_p^{\cdot}(w) = \frac{1}{\Xi(w)} \sum_N \frac{1}{N!} N(p) Q(N) w^N = \frac{1}{\Xi(w)} w_p \frac{\partial}{\partial w_p} \Xi(w)$$

as in standard probability literature.

Define $f(w)$ by $\Xi(w) = \exp(f(w))$. The quantity $f(w)$ is proportional to the *pressure* of the gas. Then

$$\frac{\partial}{\partial w_p} f(w) = \frac{1}{\Xi(w)} \frac{\partial}{\partial w_p} \Xi(w)$$

and hence $f_p^{\cdot}(w) = w_p \frac{\partial}{\partial w_p} f(w)$.

Our goal is to estimate the quantity $f_p^{\cdot}(w)$. It is well known that once this quantity is under control, then most other quantities of physical interest can be dealt with by the same method. And we will devote the following time to this aim.

At this point it is useful to introduce a physical notion that also serves to give the combinatorial interpretation that is ordinarily used by physicists.

Recall that we have a set $\mathcal{P}$, a set of *locations*. Now furthermore, we have a set U, a set of *particles*.

A *particle configuration* is a function $a : U \to \mathcal{P}$ from the finite set U to the set of locations. The interpretation is that if U has n points, then there are n particles, and the particle configuration specifies their locations.

At this point it is useful to introduce a physical notion that also serves to give the combinatorial interpretation that is ordinarily used by physicists.

Recall that we have a set $\mathcal{P}$, a set of *locations*. Now furthermore, we have a set U, a set of *particles*.

A *particle configuration* is a function $a : U \to \mathcal{P}$ from the finite set U to the set of locations. The interpretation is that if U has n points, then there are n particles, and the particle configuration specifies their locations.

There is an alternative terminology that is more common in combinatorics. The finite set U is a set of *labels*, while the set $\mathcal{P}$ is a set of *colors*. And a function $a : U \to \mathcal{P}$ is a *colored set*.

Recall the grand partition function $\Xi(w) = \sum_N \frac{1}{N!} Q(N) w^N$ introduced previously. Using some knowledge from combinatorics, it may be rewritten as an *exponential generating function* by

$$\Xi(w) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{a: U_n \to \mathcal{P}} [\prod_{\{i,j\}} (1 + t(a(i), a(j)))] \prod_i w_{a(i)},$$

taking into account of the combinatorial contribution and the product of activity variables.

Also, one may use the distributive law to write the *interaction product*, also a concept introduced before, as

$$\prod_{\{i,j\}}(1+t(a(i),a(j)))=\sum_{G}\prod_{\{i,j\}\in G}t(a(i),a(j)).$$

Here G is a *graph* on U, a set of two-element subsets $\{i,j\}$ of U. Or, in combinatorics, U is a set of *vertices*, while G is a set of *edges*.

A *connected graph* on a set U is a non-empty set of edges such that every vertex in U is connected to every other vertex in U by a *path* along edges.

It is clear that every graph on a set may be decomposed into *connected graphs* on the subsets that belong to a partition of the set.

A combinatoric fact tells us since $\Xi(w) = \exp(f(w))$, we can derive from

$$\Xi(w) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{a: U_n \to \mathcal{P}} [\prod_{\{i,j\}} (1 + t(a(i), a(j)))] \prod_i w_{a(i)}$$

to get an exponential generating function representation for $f(w)$ as

$$f(w) = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{a: U_n \to \mathcal{P}} [\sum_{G_c} \prod_{\{i,j\} \in G_c} t(a(i), a(j))] \prod_i w_{a(i)}$$

Here G_c ranges over connected graphs on the set of n particles.

Remember the expected number of particles at location p, $f_p^{\cdot}(w) = w_p \frac{\partial}{\partial w_p} f(w)$ is the ultimate quantity we are interested in. Based on the above derivations, it now has an explicit formula

$$f_p^{\cdot}(w) = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{a:U_n \to \mathcal{P}} [\sum_{\{r|a(r)=p\}} \sum_{G_c} \prod_{\{i,j\} \in G_c} t(a(i), a(j))] \prod_i w_{a(i)}$$

In combinatorial language, one is considering the contributions from rooted connected graphs with all possible choices of root of the given color (in this case, with color p).

We hope that under some suitable conditions, the power series expansion of the expected number $f_p^{\cdot}(w)$ of particles at location p converges absolutely for the given w and has absolute value bounded by some vector x pointwise.

So it's intuitive that we want to find a bound for $\sum_{G_c} \prod_{\{i,j\} \in G_c} t(a(i), a(j))$, for this purpose, we need to introduce the *Connected Graph Identity*.

We hope that under some suitable conditions, the power series expansion of the expected number $f_p^{\cdot}(w)$ of particles at location p converges absolutely for the given w and has absolute value bounded by some vector x pointwise.
So it's intuitive that we want to find a bound for $\sum_{G_c} \prod_{\{i,j\} \in G_c} t(a(i), a(j))$, for this purpose, we need to introduce the *Connected Graph Identity*.

A *tree* on a set U is a minimal connected graph on the set U. Thus every vertex may be connected to every other vertex along a unique path of edges. A *rooted tree* is a tree together with a distinguished vertex, the root.

Consider a rooted tree with a fixed root r. A tree on U may be thought of as a function $\tau : U \bigcap \{r\}^c \to U$ such that the orbit of each point ends in r. The vertices $i \neq r$ determine the edges $\{i, \tau(i)\}$ of the tree graph.

For convenience, following we will write t_{ij} for $t(a(i), a(j))$.

A *tree* on a set U is a minimal connected graph on the set U. Thus every vertex may be connected to every other vertex along a unique path of edges. A *rooted tree* is a tree together with a distinguished vertex, the root.

Consider a rooted tree with a fixed root r. A tree on U may be thought of as a function $\tau : U \bigcap \{r\}^c \to U$ such that the orbit of each point ends in r. The vertices $i \neq r$ determine the edges $\{i, \tau(i)\}$ of the tree graph.

For convenience, following we will write t_{ij} for $t(a(i), a(j))$.

A *tree* on a set U is a minimal connected graph on the set U. Thus every vertex may be connected to every other vertex along a unique path of edges. A *rooted tree* is a tree together with a distinguished vertex, the root.

Consider a rooted tree with a fixed root r. A tree on U may be thought of as a function $\tau : U \bigcap \{r\}^c \to U$ such that the orbit of each point ends in r. The vertices $i \neq r$ determine the edges $\{i, \tau(i)\}$ of the tree graph.

For convenience, following we will write t_{ij} for $t(a(i), a(j))$.

Consider a tree on a finite set U with a given root r to be a function $\tau : U \bigcap \{r\}^c \to U$. For each vertex $i \neq r$, consider an integration variable s_i in the interval $[0, 1]$ associated with the tree edge $\{i, \tau(i)\}$. Let $\mathbf{s}$ denote the family of all these variables. Let the tree factor

$$c(\tau) = \prod_{i \neq r} t_{i\tau(i)}$$

be a contribution from the edges $\{i, \tau(i)\}$ of the tree τ.

The withinlayer factor is $W(\tau) = \prod_{\{i,j\}|i\sim j \text{ mod } \tau}(1 + t_{ij})$, where the product is over two element subsets $\{i, j\}$, and $i \sim j$ mod τ means that i and j are at the same distance from the root in the tree τ.

The interlayer factor is $I(\tau, \mathbf{s}) = \prod_{i\neq r} \prod_{j|j\leftarrow i \text{ mod } \tau, j\neq\tau(i)}(1 + s_i t_{ij})$, where $j \leftarrow i$ mod τ means that j is one tree distance unit closer to the root than the tree distance of i from the root.

Then the sum over connected graphs may be expressed as a sum over trees by $\sum_{G_c} \prod_{\{i,j\}\in G_c} t_{ij} = \sum_\tau c(\tau)W(\tau) \int I(\tau, \mathbf{s})d\mathbf{s}$.

This identity looks rather impossible to understand, so let's first see what we can get from the conclusion $\sum_{G_c} \prod_{\{i,j\} \in G_c} t_{ij} = \sum_{\tau} c(\tau) W(\tau) \int I(\tau, \mathbf{s}) d\mathbf{s}$, and then go back to its proof, provided we have enough time.

An immediate corollary is the *enriched tree bound*. Remember $-1 \leq t_{ij} \leq 0$, we have

$$|\sum_{G_c} \prod_{\{i,j\} \in G_c} t_{ij}| = |\sum_{\tau} c(\tau) W(\tau) \int I(\tau, \mathbf{s}) d\mathbf{s}| \leq \sum_{\tau} c^*(\tau) Q(\tau).$$

Here the tree contribution $c(\tau)$ is estimated using absolute values by $c^*(\tau) = \prod_{i \neq r} |t_{i\tau(i)}|$.
The random interlayer contribution $I(\tau, \mathbf{s})$ is estimated by 1. Since the integration variables s_i are in the interval $[0, 1]$, $|\int I(\tau, \mathbf{s}) d\mathbf{s}| \leq 1$.

The bound for the withinlayer contribution is more subtle. Recall it is defined by $W(\tau) = \prod_{\{i,j\}|i\sim j \text{ mod } \tau}(1 + t_{ij})$, we now only consider the *fiber contribution* given by

$$Q(\tau) = \prod_{\{i,j\}|i\downarrow j \text{ mod } \tau} (1 + t_{ij})$$

Here $i \downarrow j$ mod τ meaning that $\tau(i) = \tau(j)$. This is the contribution from edges $\{i, j\}$ that belong to the fiber over a vertex. Thus only a small part of the withinlayer interaction in the connected graph is retained.

The above considerations give us a bound $g_p(w)$ for $f_p^{\cdot}(w)$, given by

$$\sum_{n=1}^{\infty} \frac{1}{n!} \sum_{a:U_n \to \mathcal{P}} [\sum_{\{r|a(r)=p\}} \sum_{T_r} \prod_{\{i,j\} \in T_r} |t_{ij}| \prod_{i \downarrow j} (1 + t_{ij})] \prod_i w_{a(i)}$$

Here T_r is a rooted tree with root r in U_n. The task for us now is to find a way to deal with this rather complicated sum and show that $g_p(w)$ converges. We will concentrate on this in the following.

Such a rooted tree is defined recursively. The main ingredient is a construction that describes the edges from an external root of color p to points of various colors q and with the interaction among these points. And the corresponding exponential generating function is

$$\Xi_p(w) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{a:U_n \to \mathcal{P}} \prod_j |t(p, a(j))| \prod_{\{i,j\}} (1 + t(a(i), a(j))) \prod_j w_{a(j)}$$

$$= \sum_K \frac{1}{K!} |t_p|^K Q(K) w^K.$$

There is a systematic procedure for going from combinatoric constructions to corresponding operations on exponential generating functions.

Say that we have a set W and want to partition it into subsets and have a rooted subtree on each subset. Furthermore, we want to associate with each subset in the partition the color coming from the root of the corresponding rooted subtree. Finally, we want to have an interaction of such roots with each other and with an external point of color p. This combined construction corresponds to the composition $\Xi_p(g(w))$.

There is a systematic procedure for going from combinatoric constructions to corresponding operations on exponential generating functions.

Say that we have a set W and want to partition it into subsets and have a rooted subtree on each subset. Furthermore, we want to associate with each subset in the partition the color coming from the root of the corresponding rooted subtree. Finally, we want to have an interaction of such roots with each other and with an external point of color p. This combined construction corresponds to the composition $\Xi_p(g(w))$.

The other ingredient is the construction that picks a single point of color p, designed to be the root. This corresponds to the exponential generating function that is just a coordinate function w_p.

Finally, we want to make this all into one tree on the set U by adding a root of color p and doing the above construction on the complement W. And the result is the exponential generating function $w_p \Xi_p(g(w))$.

We recognize this is the recursive construction of the original tree. The conclusion is that $g(w)$ satisfies the fixed point equation $g_p(w) = w_p \Xi_p(g(w))$.

The other ingredient is the construction that picks a single point of color p, designed to be the root. This corresponds to the exponential generating function that is just a coordinate function w_p.

Finally, we want to make this all into one tree on the set U by adding a root of color p and doing the above construction on the complement W. And the result is the exponential generating function $w_p \Xi_p(g(w))$.

We recognize this is the recursive construction of the original tree. The conclusion is that $g(w)$ satisfies the fixed point equation $g_p(w) = w_p \Xi_p(g(w))$.

The other ingredient is the construction that picks a single point of color p, designed to be the root. This corresponds to the exponential generating function that is just a coordinate function w_p.

Finally, we want to make this all into one tree on the set U by adding a root of color p and doing the above construction on the complement W. And the result is the exponential generating function $w_p \Xi_p(g(w))$.

We recognize this is the recursive construction of the original tree.

The conclusion is that $g(w)$ satisfies the fixed point equation $g_p(w) = w_p \Xi_p(g(w))$.

Define $\phi(x)_p = w_p \Xi_p(x)$. The previous recursive equation for the exponential generating function is then the fixed point equation $\phi(z) = z$. What we need to do is to show the existence of a fixed point and also show that it is finite.

For this purpose, we need a quick review of the *Knaster-Tarski Theorem*. Assuming ϕ as defined above is increasing.

An increasing function ϕ from a complete lattice to itself has a fixed point. And the least fixed point is given by $z = \inf\{x | \phi(x) \leq x\}$.

Proof.

Let S be the set of x with $\phi(x) \leq x$. Since the lattice is complete, S has an infimum z.
Consider arbitrary x in S. We have $z \leq x$. As we are given ϕ increasing, $\phi(z) \leq \phi(x) \leq x$. Thus $\phi(z)$ is a lower bound for S. As such, $\phi(z) \leq z$.
Using ϕ increasing again, $\phi(\phi(z)) \leq \phi(z)$, from which we know $\phi(z)$ is in S. Thus $z \leq \phi(z)$.
Conclusion: $\phi(z) = z$.

□

An increasing function ϕ from a complete lattice to itself has a fixed point. And the least fixed point is given by $z = \inf\{x | \phi(x) \leq x\}$.

Proof.

Let S be the set of x with $\phi(x) \leq x$. Since the lattice is complete, S has an infimum z.
Consider arbitrary x in S. We have $z \leq x$. As we are given ϕ increasing, $\phi(z) \leq \phi(x) \leq x$. Thus $\phi(z)$ is a lower bound for S. As such, $\phi(z) \leq z$.
Using ϕ increasing again, $\phi(\phi(z)) \leq \phi(z)$, from which we know $\phi(z)$ is in S. Thus $z \leq \phi(z)$.
Conclusion: $\phi(z) = z$.

□

It is also useful to know an iterative way to get the least fixed point z.
The obvious strategy is to let $u^{\{0\}}$ be the least element of the complete lattice, and define the sequence $u^{\{n+1\}} = \phi(u^{\{n\}})$.
Since ϕ is increasing, it follows by induction that the sequence $u^{\{n\}}$ is increasing. It also follows by induction that each $u^{\{n\}} \leq z$, where z is the least fixed point. Let $z' = \sup_n u^{\{n\}}$, then $z' \leq z$.
Assume the monotone convergence property for increasing sequences, which holds in general applications, $\sup_n \phi(u^{\{n\}}) = \phi(\sup_n u^{\{n\}})$. This says that $z' = \sup_n u^{\{n+1\}} = \phi(z')$, thus $z' = z$ is indeed the fixed point z.

Here is the Fernández-Procacci result.

Theorem

Consider the discrete gas model with interaction factor $t(p,q)$ satisfying $-1 \leq t(p,q) \leq 0$, and take the activities $w_p \geq 0$. Suppose that there exists a finite vector $x \geq 0$ such that

$$w_p \sum_K \frac{1}{K!} |t_p|^K Q(K) x^K \leq x_p.$$

Then the power series expansion of the expected number $f_p^{\cdot}(w)$ of particles at location p converges absolutely for the given w and has absolute value bounded by x_p.

It should not be too hard to understand this theorem now.
In the present context the natural choice of complete lattice is $[0, +\infty]^{\mathcal{P}}$. The condition in the Fernández-Procacci result is just $\phi(x) \leq x$ with x finite, using our previous notation. This guarantees that the least fixed point $g_p(w) = \inf\{x | \phi(x) \leq x\}$ is itself finite, with $g_p(w) \leq x$.
Finally, recall that the fixed point $g_p(w)$ is an absolute value bound for $f_p^{\cdot}(w)$. Its finiteness and convergence will imply the absolute convergence of the expected number $f_p^{\cdot}(w)$.

It should not be too hard to understand this theorem now.
In the present context the natural choice of complete lattice is $[0, +\infty]^{\mathcal{P}}$. The condition in the Fernández-Procacci result is just $\phi(x) \leq x$ with x finite, using our previous notation. This guarantees that the least fixed point $g_p(w) = \inf\{x | \phi(x) \leq x\}$ is itself finite, with $g_p(w) \leq x$.
Finally, recall that the fixed point $g_p(w)$ is an absolute value bound for $f_p^{\cdot}(w)$. Its finiteness and convergence will imply the absolute convergence of the expected number $f_p^{\cdot}(w)$.

As I have promised, if we have enough time, we will trace through the proof of the Connected Graph Identity.
Maybe it's better to look at this identity again first.

Consider a tree on a finite set U with a given root r to be a function $\tau : U \bigcap \{r\}^c \to U$. For each vertex $i \neq r$, consider an integration variable s_i in the interval $[0, 1]$ associated with the tree edge $\{i, \tau(i)\}$. Let $\mathbf{s}$ denote the family of all these variables. Let the tree factor

$$c(\tau) = \prod_{i \neq r} t_{i\tau(i)}$$

be a contribution from the edges $\{i, \tau(i)\}$ of the tree τ.

The withinlayer factor is $W(\tau) = \prod_{\{i,j\}|i\sim j \text{ mod } \tau}(1 + t_{ij})$, where the product is over two element subsets $\{i, j\}$, and $i \sim j$ mod τ means that i and j are at the same distance from the root in the tree τ.

The interlayer factor is $I(\tau, \mathbf{s}) = \prod_{i\neq r} \prod_{j|j\leftarrow i \text{ mod } \tau, j\neq\tau(i)}(1 + s_i t_{ij})$, where $j \leftarrow i$ mod τ means that j is one tree distance unit closer to the root than the tree distance of i from the root.

Then the sum over connected graphs may be expressed as a sum over trees by $\sum_{G_c} \prod_{\{i,j\}\in G_c} t_{ij} = \sum_\tau c(\tau)W(\tau) \int I(\tau, \mathbf{s})d\mathbf{s}$.

The first thing to do is to rewrite $\sum_{G_c} \prod_{\{i,j\} \in G_c} t_{ij}$.
Fix a root r in U. Consider a decomposition $\Delta = (U_0, U_1, ..., U_h)$ of U into non-empty subsets such that $U_0 = \{r\}$.
Let $\mathcal{G}_c(\Delta)$ consist of all connected graphs G_c with the property that the set of points in G_c a graph distance m from the root is U_m.
We then have

$$\sum_{G_c} \prod_{\{i,j\} \in G_c} t_{ij} = \sum_{\Delta} \sum_{G_c \in \mathcal{G}_c(\Delta)} \prod_{\{i,j\} \in G_c} t_{ij}.$$

Not-so-detailed Proof of the Connected Graph Identity

There is a combinatoric fact that $\sum_{G_c \in \mathcal{G}_c(\Delta)} \prod_{\{i,j\} \in G_c} t_{ij} = W_\Delta I_\Delta$.
Here the withinlayer factor is $W_\Delta = \prod_{i \sim j}(1 + t_{ij})$.
And the interlayer factor is $I_\Delta = \prod_{i \neq r}[\prod_{j \leftarrow i}(1 + t_{ij}) - 1]$.

Also, we observe that

$$\prod_{j\leftarrow i}(1+t_{ij})-1=\int_0^1\frac{d}{ds}\prod_{j\leftarrow i}(1+st_{ij})ds=\sum_{k\leftarrow i}[\int_0^1 t_{ik}\prod_{j\leftarrow i,j\neq k}(1+st_{ij})ds].$$

But notice that

$$I_\Delta=\prod_{i\neq r}[\sum_{k\leftarrow i}\int_0^1 t_{ik}\prod_{j\leftarrow i,j\neq k}(1+st_{ij})ds]$$

$$=\sum_\tau\prod_{i\neq r}[\int_0^1 t_{i\tau(i)}\prod_{j\leftarrow i,j\neq\tau(i)}(1+st_{ij})ds]$$

Denote the set of all tree functions τ by $\mathcal{T}(\Delta)$.

We can write the product of integrals of I_Δ as a multiple integral of a product. And we have

$$I_\Delta = \sum_{\tau \in \mathcal{T}(\Delta)} \int_0^1 \cdots \int_0^1 \prod_{i \neq r} [t_{i\tau(i)} \prod_{j \leftarrow i, j \neq \tau(i)} (1 + s_i t_{ij}) ds_i].$$

Since the decomposition is determined by the tree distance, we can write this equivalently as

$$I_\Delta = \sum_{\tau \in \mathcal{T}(\Delta)} c(\tau) \int I(\tau, \mathbf{s}) d\mathbf{s}$$

And since the withinlayer factor also depends on the decomposition via the tree, we may write W_Δ as $W(\tau)$. And we obtain

$$\sum_{G_c \in \mathcal{G}_c(\Delta)} \prod_{\{i,j\} \in G_c} t_{ij} = \sum_{\tau \in \mathcal{T}(\Delta)} c(\tau) W(\tau) \int I(\tau, \mathbf{s}) d\mathbf{s}.$$

Summing over Δ will complete the proof of the connected graph identity.

Thank You!