Thermodynamic Binding Networks Capture Equilibrium Monomer-Polymer Thermodynamics

And Kaan Ata Yilmaz
Electrical and Computer Engineering, University of Texas at Austin, TX, USA

David Soloveichik
Electrical and Computer Engineering, University of Texas at Austin, TX, USA

Austin Luchsinger
Computer Science, University of Texas Rio Grande Valley, Edinburg, TX, USA

Abstract

Thermodynamic Binding Networks (TBNs) are a minimal model for molecular thermodynamics. Despite their simplicity, the stable behavior of these systems can be unexpectedly nontrivial. To better understand the expressive power of the model, we compare TBNs with Thermodynamic Affinity Networks (TANs), a partition-scoring formalism in which polymer-level favorability is specified directly, rather than derived from domain-level interactions. We show that these two models are thermodynamically equivalent at the level of stable configurations: every TBN can be represented directly by a TAN, and every TAN can be indirectly realized by a TBN whose stable configurations map exactly to the stable configurations of the TAN. The key idea in the latter direction is to use auxiliary monomers that allow the system to convert polymer-level free-energy rewards specified by the TAN into entropic gains from separate complexes in the TBN. This correspondence shows that the local rules of TBNs do not limit their equilibrium expressive power nearly as much as one might expect. In this minimum-free-energy sense of equilibrium, TBNs capture the coarse-grained monomer-polymer thermodynamics formalized by TANs.

2012 ACM Subject Classification Theory of computation → Models of computation

Keywords and phrases Thermodynamic Binding Networks, Thermodynamic Affinity Networks, Equilibrium Thermodynamics

Digital Object Identifier 10.4230/LIPIcs.DNA.32.9

Funding Schmidt Sciences Polymath Award to D.S., DOE grant DE-SC0024248, NSF SemiSynBio III: GOALI grant 2227578.

1 Introduction

Thermodynamic Binding Networks (TBNs) were introduced as a deliberately simplified model of molecular thermodynamics [7]. In this model, monomers are multisets of domain types, polymers are assemblies of monomers (which may or may not be joined by pairwise complementary domain bindings), and configurations are collections of polymers. Each configuration is assigned an energy based on the number of bonds (enthalpy) and the number of polymers it contains (entropy), and the configurations with the lowest (most favorable) energy are said to be stable. See Figure 1 for an example TBN system.

At first glance, the expressive power of TBNs seems limited. The model is intentionally simplified: it lacks several mechanisms that have proved powerful in other models, including geometry [9, 19], non-complementary domain interactions [1, 12], and higher-order regulatory effects such as allostery [13, 16]. One might therefore expect TBNs to admit only a narrow and fairly transparent class of thermodynamic behaviors. However, despite this simplicity, TBNs have proven surprisingly difficult to analyze, and establishing formal properties of their behavior is often quite subtle [3, 4, 10, 14]. This raises a natural question: to what extent can TBNs capture arbitrary monomer-polymer thermodynamics?

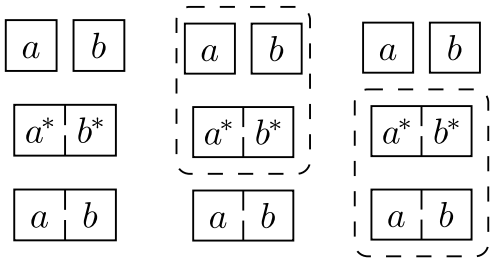
To make this question more precise, we compare TBNs with a more abstract thermodynamic formalism: Thermodynamic Affinity Networks (TANs). Equilibrium chemical thermodynamics is often understood at the level of macrostates that record only coarse-grained information about the system, such as which molecular complexes are present and in what counts, rather than the full microscopic arrangement of the molecules. The thermodynamic favorability of such a macrostate is determined by the total statistical weight of the microstates that realize it. This statistical weight combines energetic effects (such as favorable molecular interactions within complexes) with entropic effects (such as translational freedom and combinatorial multiplicity). TANs formalize this coarse-grained macrostate viewpoint for finite monomer-polymer systems.

A TAN consists of a finite multiset of monomer types M along with an affinity function $\mathcal{F} : \mathcal{P}(M) \to \mathbb{Q}_{\geq 0}$ that assigns a nonnegative value to every possible polymer. A TAN configuration is simply a partition of the monomers into polymers, and a configuration's thermodynamic preference is determined by the total affinity of those polymers as well as the number of polymers in the partition. This abstraction removes the domain-level mechanics entirely and instead treats polymer favorability as a primitive. TANs should be viewed as a finite formal language for monomer-polymer macrostate preferences rather than as a new model for chemical thermodynamics. Mathematically, TANs are closely related to other partition-scoring formalisms such as combinatorial auctions and coalition structure generation [5, 15].

The contrast between TBNs and TANs becomes apparent even in the simplest systems. As an example, consider a TAN system with three monomers m_1, m_2, and m_3 where the polymer $\{m_1, m_2, m_3\}$ has positive affinity (i.e., $\mathcal{F}(\{m_1, m_2, m_3\}) > 0$), and all pairwise combinations of the monomers have zero affinity. This kind of polymer-level preference is natural in the TAN setting; however, it is not clear how to achieve this behavior in a TBN since the enthalpic contribution of a polymer comes from the pairwise complementary bonds within it. In that setting, any bond present in the ternary complex would already be a bond between some pair of monomers and would therefore create pairwise affinity.

Although a *direct* TBN implementation of the previous example is not possible, the main result of this work shows that TBNs can *indirectly* capture this behavior, as well as any other behavior described by an arbitrary affinity function. To accomplish this, our key technique takes the favorability assigned directly to polymers in a TAN and encodes it indirectly in a TBN through entropy and particle counts, rather than explicit bonds. The idea is that a TBN realization of this example consists of three large complexes representing the TAN monomers. When those complexes merge to form a single polymer, they release a precise number of auxiliary "helper" monomers. The construction is engineered so that the entropy increase from these released monomers lowers the free energy of the TBN configuration by an amount that is consistent with the TAN configuration (e.g., accounting for the affinity of the ternary complex).

Ultimately, we show that every TAN can be represented by a thermodynamically equivalent TBN, and conversely that every TBN has an equivalent TAN. This surprising equivalence shows that TANs do not extend the range of thermodynamic behaviors realizable by TBNs; rather, they provide a more direct language for specifying those behaviors. Since TANs formalize a highly general class of coarse-grained monomer-polymer thermodynamics, this equivalence suggests that TBNs capture this class as well. We note that throughout the paper, "equilibrium" is meant in the minimum-free-energy sense: our realization definition compares the stable configurations of each system, rather than full Boltzmann equilibrium distributions. With this convention, and given the natural implementation of TBNs with DNA strands, our result supports the view that DNA-based binding is a broadly expressive substrate for TAN-style monomer-polymer thermodynamics.



(a) TBN $\mathcal{B}$. **(b)** config. β. **(c)** config. δ.

Figure 1 (a) An example TBN $\mathcal{B}$ with $T = \{\{a\}, \{b\}, \{a, b\}, \{a^*, b^*\}\}$, (b) a saturated configuration β of $\mathcal{B}$, and (c) a stable configuration δ of $\mathcal{B}$.

Paper Organization

In Section 2, we introduce the necessary definitions of TBNs, TANs, and our notions of thermodynamic realization and equivalence. In Section 3, we present an initial result establishing that any given TBN can be directly realized by a TAN. Section 4 introduces a generalization of the grid gate from [3], which serves as a key module in our construction of a TBN that is thermodynamically equivalent to a given TAN. In Section 5, we present the main construction of this work: a TBN that thermodynamically realizes a given TAN in the large-volume regime. Section 6 extends the construction to the general TAN energy function by incorporating combinatorial entropy. Finally, Section 7 concludes with a discussion of implications and directions for future work.

2 Preliminaries

We first provide definitions for the two models, Thermodynamic Affinity Networks and Thermodynamic Binding Networks, and then provide our definition of thermodynamic equivalence.

2.1 Thermodynamic Binding Network

A *Thermodynamic Binding Network* $\mathcal{B} = (D, T)$ is a tuple where:

- D is a finite set of domain types,
- T is a finite multiset of monomer types over $D \cup D^*$

A *domain type* is a formal symbol. D^* denotes the set of complementary domain types for D. We often refer to domain types from D as unstarred domains and domain types from D^* as starred domains. A *monomer type* is a multiset of domain types. We call an instance of a domain type a *domain* and an instance of a monomer type a *monomer*. A *configuration* β of a TBN is a partition of T into parts, which we call *polymers*. A *bond* is a pairwise matching of complementary domains within a polymer. Note that a polymer is simply a block of the configuration partition, and has no requirement on the formation of bonds between its constituent monomers. Furthermore, bonds are not explicitly specified in a configuration; we just intuitively think of polymers as having a maximal matching[1].

[1] We use the term matching in the graph theory sense. Indeed, if we created a graph of the polymer using domains as vertices with edges between complementary domains, a matching of that graph would yield a set of bonds in the polymer.

The *entropy* $S(\beta)$ of configuration β is the number of polymers in β. The *enthalpy* $H(\beta)$ of configuration β is the number of bonds in β. We use the convention that TBNs are *star-limiting*: for each domain type, there are always at least as many instances of the unstarred type as the starred type among all monomers. A starred (unstarred) domain within a polymer is said to be *exposed* if, among the monomers in that polymer, the number of instances of the domain type exceeds the number of unstarred (starred) instances of the complementary domain type. A polymer is *self-saturated* if it contains no exposed starred domains. A self-saturated polymer is *splittable*, if it can be split into two self-saturated polymers and *unsplittable*, otherwise. For a star-limiting TBN, a configuration is *saturated* if every polymer in the configuration is self-saturated. Intuitively, this means that no polymer contains an unmatched starred domain and equivalently, all possible bonds are formed.

The *energy* of β is $E(\beta) = -wH(\beta) - S(\beta)$ where w is an enthalpy weighting parameter. The configurations that minimize E are said to be *stable*. As is common when discussing TBNs, we operate under a regime where w is high enough for enthalpy to be considered infinitely more favorable than entropy. The following lemma identifies a key structural constraint of stable configurations for $w > 1$ and shows that stability is insensitive to the exact value of w.

▶ **Lemma 1.** *For a given TBN $\mathcal{B}$, if $w > 1$, a stable configuration of $\mathcal{B}$ is an entropy maximizing saturated configuration. Furthermore, the set of stable configurations of $\mathcal{B}$ is the same for any $w > 1$.*

Proof. Suppose β is not a saturated configuration. Then there exist two polymers P_1 and P_2 in β such that P_1 contains an exposed starred domain, without loss of generality a^*, and P_2 contains an exposed unstarred domain a. Consider a configuration β' obtained from β by replacing P_1 and P_2 with their union $P_1 \uplus P_2$, while leaving all other polymers unchanged. In this case, we have $H(\beta') \geq H(\beta) + 1$ and $S(\beta') = S(\beta) - 1$. Therefore, $E(\beta') - E(\beta) \leq -(w-1)$. As $w > 1$, β cannot be stable. Consequently, for $w > 1$, every stable configuration is saturated, i.e., it realizes all possible bonds.

Among the saturated configurations, stability further favors configurations that maximize entropy, namely those with the largest number of free complexes. Moreover, since all saturated configurations achieve the same maximum enthalpy (i.e., number of bonds), which is fixed for a given TBN, the set of stable configurations is identical for all values of $w > 1$. ◀

Note that if there is a splittable polymer in a saturated configuration, the entropy of the configuration can be increased by 1 by splitting it. Hence, in a stable configuration of a TBN, there cannot be any splittable polymers.

2.2 Thermodynamic Affinity Network

The TAN model provides a more abstract formalism for thermodynamic networks.

A *Thermodynamic Affinity Network* $\mathcal{A} = (M, \mathcal{F}, u, v)$ is a 4-tuple consisting of:

- M, a finite multiset of monomer types,
- $\mathcal{F}$, an affinity function (defined below),
- $u, v \in \mathbb{Q}_{\geq 0}$, enthalpy and entropy weight parameters.[2]

A *monomer type* is a formal symbol, and a *polymer type* is a multiset of monomer types. We define $\mathcal{P}(M) = \{N \mid N \text{ is a nonempty sub-multiset of } M\}$, i.e., the set of all polymer types over M. The affinity function $\mathcal{F} : \mathcal{P}(M) \to \mathbb{Q}_{\geq 0}$ assigns to each polymer

[2] Intuitively, one may view $\frac{v}{u}$ as the temperature of the system.

its enthalpic contribution upon formation. As in the TBN setting, a *configuration* α of a TAN is a partition of M, and may be viewed as a multiset of polymers. We write $\text{supp}(\alpha) = \{P_1, P_2, \ldots, P_n\}$ for the set of distinct polymers appearing in α, and let $\underline{\alpha}(s)$ denote the multiplicity of $P_s \in \text{supp}(\alpha)$. Throughout the paper, we alternate between indexing directly over α (treating identical polymers as distinct copies, as in Section 5) and indexing over $\text{supp}(\alpha)$ (as here and in Section 6).

The energy of a TAN configuration α is defined by

$$E_{\text{TAN}}(\alpha) = -\sum_{P_s \in \text{supp}(\alpha)} \underline{\alpha}(s)\, w \log(V)\, \mathcal{F}(P_s) - \sum_{P_s \in \text{supp}(\alpha)} \underline{\alpha}(s) \log(V) + \sum_{P_s \in \text{supp}(\alpha)} \sum_{j=1}^{\underline{\alpha}(s)} \log j,$$

where V is the volume and $w \in \mathbb{Q}$ is a constant. A derivation from first principles is provided in Appendix A. We approximate $\log V$ to five digits, making it rational, and define $u = w \log V$ and $v = \log V$, so that $u, v \in \mathbb{Q}$. The entropy $S(\alpha)$ is the number of polymers in α, and the enthalpy $H(\alpha)$ is $\sum_{P_s \in \text{supp}(\alpha)} \underline{\alpha}(s)\mathcal{F}(P_s)$. With these definitions, the energy can be written compactly as

$$E(\alpha) = -uH(\alpha) - vS(\alpha) + \sum_{P_s \in \text{supp}(\alpha)} \sum_{j=1}^{\underline{\alpha}(s)} \log j.$$

The final term, arising from indistinguishability of polymers (see Appendix A), is referred to as the *combinatorial entropy*. A configuration α is *stable* if it minimizes $E(\alpha)$. Notice that, since the TAN model has no notion of bonds, there is no need to define saturated configurations.

2.3 Configuration Mappings and Realizations

To compare the expressive power of these thermodynamic models, we relate their state spaces via mappings. Intuitively, such mappings interpret configurations of one system as configurations of another, allowing us to formalize when two systems exhibit equivalent behavior. Two notions of equivalence are natural in this setting. The first is purely thermodynamic: under a given mapping, two systems have equivalent equilibria (i.e., roughly the same stable configurations). The second is kinetic: under a given mapping, two systems have equivalent dynamics[3]. In this work, we focus only on thermodynamic equivalence.

Here we define what it means for a Thermodynamic Binding Network (TBN) to be thermodynamically equivalent to a Thermodynamic Affinity Network (TAN) by way of configuration mappings that we refer to as realizations. Throughout this subsection, let X denote the *realizing* system and Y denote the *realized* system. The definitions in this subsection are model-independent: both the realizing system and the realized system may be either a TAN or a TBN.

Mappings

Our main object of interest is a configuration-level map (that we call R_{conf}) which interprets a configuration of the realizing system X as a corresponding configuration of the realized system Y. Rather than allowing an arbitrary mapping between configurations, we ensure R_{conf} is "reasonable" by building up from more fundamental mappings between monomer

[3] This kinetic equivalence is well studied as simulation/bisimulation [2, 8, 11]

types (R_{type}) and polymers (R_{poly}). Informally, R_{type} specifies which monomer types of Y (if any) each monomer type from X corresponds to. R_{poly} naturally extends R_{type} to map polymers, and R_{conf} extends R_{poly} in the same way.

Both models have a finite multiset of monomer types and configurations are given by partitions of those multisets into polymers. Let M_X and M_Y denote the multisets of monomer types of X and Y, respectively, and let S_X and S_Y denote their underlying sets of monomer types. Let $\texttt{Conf}(X)$ and $\texttt{Conf}(Y)$ denote the sets of configurations of X and Y, and let $\texttt{Stab}(X) \subseteq \texttt{Conf}(X)$ and $\texttt{Stab}(Y) \subseteq \texttt{Conf}(Y)$ denote their stable configurations.

A *type map* is a partial function $R_{\text{type}} : S_X \to S_Y$ that maps monomer types from system X to system Y. We say R_{type} is *valid* if for every monomer type $y \in S_Y$, there exists exactly one monomer type $x \in S_X$ such that $R_{\text{type}}(x) = y$ and the multiplicity of x in M_X equals the multiplicity of y in M_Y. If $R_{\text{type}}(x)$ is undefined, we say $R_{\text{type}}(x) = \emptyset$. Unmapped monomer types represent auxiliary components of the realizing system X that do not correspond to monomers of the realized system Y. We can naturally extend R_{type} to map polymers and configurations as well. With P denoting a multiset of monomer types (i.e., a polymer) of X and for a valid R_{type}, define $R_{\text{poly}}(P) = \{\!|R_{\text{type}}(x) \mid x \in P \text{ and } x \in dom(R_{\text{type}})|\!\}$. And similarly, with α denoting a multiset of polymers (i.e., a configuration) of X, define $R_{\text{conf}}(\alpha) = \{\!|R_{\text{poly}}(P) \mid P \in \alpha \text{ and } R_{\text{poly}}(P) \neq \emptyset|\!\}$. In other words, R_{poly} ignores unmapped monomers and R_{conf} ignores polymers of unmapped monomers. Note that multiple configurations of X may be mapped by R_{conf} to the same configuration of Y under this definition.

Thermodynamic Realizations

We say system X *thermodynamically realizes* system Y under a valid type map R_{type} if the following conditions hold:

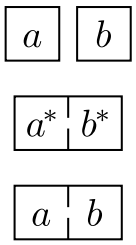
1. For every $\alpha \in \texttt{Conf}(Y)$, there exists a $\beta \in \texttt{Conf}(X)$ such that $R_{\text{conf}}(\beta) = \alpha$,
2. if $\omega \in \texttt{Stab}(X)$ then $R_{\text{conf}}(\omega) \in \texttt{Stab}(Y)$, and
3. for every $\rho \in \texttt{Stab}(Y)$, there exists a $\varphi \in \texttt{Stab}(X)$ such that $R_{\text{conf}}(\varphi) = \rho$.

The first condition says that every possible configuration of the realized system has at least one representation in the realizing system. The second says that stable configurations of the realizing system must map to stable configurations of the realized system. The third says that every stable configuration of the realized system is represented by at least one stable configuration of the realizing system.

A valid realization is a *direct realization* if R_{type} is total (i.e., every monomer type of the realizing system maps to a monomer type of the realized system), and it is an *indirect realization* otherwise. Thus, indirect realizations permit extra "helper" monomer types in the realizing system.

We say two systems are *thermodynamically equivalent (under realization)* if each one thermodynamically realizes the other, possibly under different realization maps. Again, this notion is completely based on comparing the stable configurations of the systems under a configuration map. It does not, by itself, require the two systems to have the same kinetic pathways, transition barriers, or local minima.

In Section 3, we instantiate these definitions with a TAN as the realizing system and a TBN as the realized system. In Sections 5 and 6 we reverse this direction and consider a TBN as the realizing system and a TAN as the realized system.



$$M = \{m_1, m_2, m_3, m_4\};\ \ u = 2, v = 1$$
$$\mathcal{F}(\{m_1, m_4\}) = 1$$
$$\mathcal{F}(\{m_2, m_4\}) = 1$$
$$\mathcal{F}(\{m_3, m_4\}) = 2$$
$$\mathcal{F}(\{m_1, m_3, m_4\}) = 2$$
$$\mathcal{F}(\{m_2, m_3, m_4\}) = 2$$
$$\mathcal{F}(\{m_1, m_2, m_4\}) = 2$$
$$\mathcal{F}(\{m_1, m_2, m_3, m_4\}) = 2$$

Figure 2 An explicit example of a TBN $\mathcal{B}$ (left) and a TAN $\mathcal{A}$ (right) that thermodynamically realizes it. For brevity, all instances where $\mathcal{F}(p) = 0$ are omitted.

3 Realizing TBNs with TANs

The first (rather straightforward) result is that for any TBN, there exists a TAN that directly realizes it. Figure 2 shows an explicit example of a TAN that thermodynamically realizes a TBN.

▶ **Lemma 2.** *For any TBN $\mathcal{B}$, there exists a TAN $\mathcal{A}$ such that $\mathcal{A}$ directly realizes $\mathcal{B}$.*

Proof. Given $\mathcal{B} = (D, T)$, construct $\mathcal{A} = (M, \mathcal{F}, u, v)$ as follows: For every monomer type $t \in T$, create a monomer type $m \in M$ and let $R_{\text{type}}(m) = t$, and set the multiplicity of each monomer type m to match that of $R_{\text{type}}(m) = t$. This implies that the first condition of thermodynamic realization is satisfied. Furthermore, since there are no unmapped monomers in $\mathcal{A}$, this is a direct realization and it follows that for any TBN configuration β there is exactly one TAN configuration α such that $R_{\text{conf}}(\alpha) = \beta$. As R_{conf} is defined for all TAN configurations, this implies that R_{conf} is a bijection. Now, for every possible polymer type P in $\mathcal{P}(M)$, let $\mathcal{F}(P)$ be the number of bonds in $R_{\text{poly}}(P)$ if $R_{\text{poly}}(P) \neq \emptyset$, and 0 otherwise. By Lemma 1, a stable configuration of $\mathcal{B}$ minimizes the energy function $E_{\text{TBN}}(\beta') = -wH_{\text{TBN}}(\beta') - S(\beta')$, for any value $w > 1$. Then, it suffices to set $u = 2C_{\text{Vol}}$ and $v = C_{\text{Vol}}$ (effectively $w = 2$), where $C_{\text{Vol}} = \log V$ is sufficiently large such that the combinatorial entropy term is negligible. The choice of $\mathcal{F}$ implies $H_{\text{TAN}}(\alpha) = H_{\text{TBN}}(R_{\text{conf}}(\alpha))$. Moreover, $S(\alpha) = S(R_{\text{conf}}(\alpha))$. Then $E_{\text{TAN}}(\alpha) = C_{\text{Vol}}E_{\text{TBN}}(R_{\text{conf}}(\alpha))$. As R_{conf} is a bijection, α minimizes E_{TAN} if and only if $R_{\text{conf}}(\alpha)$ minimizes E_{TBN}. Hence, $\mathcal{A}$ directly realizes $\mathcal{B}$. ◀

4 Rectangular Grid Gates

The work of [3] introduced a very useful module known as the grid gate. Our construction in Section 5 will employ a generalized version of this module that is rectangular, rather than square (Figure 3). Following the definition of square grid gates in [3], the rectangular grid gate B_{hw} with parameters h (for height) and w (for width) includes the following monomer types: *"horizontal"* $\mathcal{H}_i = (x_i)^w$ for $i \in \{1, 2, ..., h\}$, where $(x_i)^w$ denotes w-many copies of the binding site x_i, *"grid base"* $G = \{(x_i^*)^w\}_{i=1}^h$, and w-many copies of *"verticals"* $\mathcal{V} = \{x_i\}_{i=1}^h$. Figure 3 shows an example of a rectangular grid gate B_{35}. In the figure, we use colors to indicate domains and the star sign to indicate the starred versions of those domains.

Note that both the set of horizontal monomers, $\{\mathcal{H}_i\}_{i=1}^h$, and the w copies of vertical monomers, $(\mathcal{V})^w$, can saturate the grid base individually. As each pair of horizontal and vertical monomers of a rectangular grid gate shares exactly one domain, as in the square grid gate, we can immediately generalize Lemma 24 of [3] as:

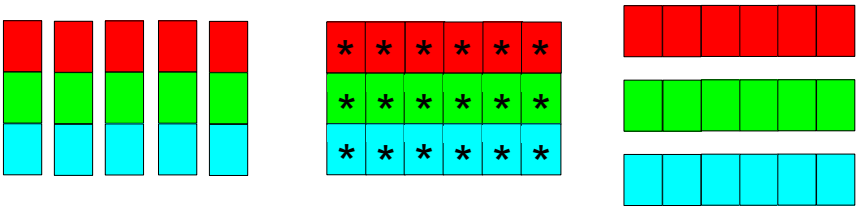


Figure 3 An example of a rectangular grid gate with parameters $h = 3$, $w = 5$.

▶ **Lemma 3.** *In a saturated configuration of B_{hw}, a polymer containing the grid base G also contains $\{\mathcal{H}_i\}_{i=1}^{h}$ or $(\mathcal{V})^w$ as a subset.*

5 Realizing TANs with TBNs: Large-Volume Regime

In Sections 5 and 6, we present the main constructions of this paper, showing how to construct a TBN $\mathcal{B}$ that indirectly realizes a given TAN $\mathcal{A} = (M, \mathcal{F}, u, v)$. We first consider the regime in which the volume parameter V is sufficiently large so that the combinatorial entropy contribution becomes negligible. In this case, the TAN energy function reduces to $E_{\text{TAN}}(\alpha) = -uH(\alpha) - vS(\alpha)$. Without loss of generality, we assume $u, v \in \mathbb{N}$ and $\mathcal{F} : \mathcal{P}(M) \to \mathbb{Z}_{\geq 0}$. The construction presented in Section 6 for the general energy function setting is built directly upon the construction introduced in this section. For the remainder of this section, we fix $k = u \max_P \mathcal{F}(P) + 2|M| + v + 1$.

Construction sketch: At a high level, polymers of $\mathcal{A}$ are managed by rectangular grid gates in $\mathcal{B}$ whose horizontal monomers map to monomers in $\mathcal{A}$. Rectangular grid gates in the vertical configuration map to nothing in the TAN. Rectangular grid gates in the horizontal configuration represent the formation of that polymer in the TAN. The number of vertical monomers freed by placing a rectangular grid gate in the horizontal configuration corresponds to the energy reward for forming that polymer in the TAN. The formal construction is detailed below.

Mapped Monomers. For each monomer type m_i in $\mathcal{A}$, we introduce a distinct domain type in $\mathcal{B}$ associated with m_i, which we denote without loss of generality by m_i, and define a corresponding monomer type consisting of k copies of this domain. The multiplicity of each such monomer type in $\mathcal{B}$ is set to match the multiplicity of m_i in $\mathcal{A}$. We refer to these as *mapped monomers*, since the domain of R_{type} consists precisely of their types. Note that R_{type} is valid and the first condition of thermodynamic realization is satisfied. We refer to a mapped monomer consisting of k copies of domain m_i as a *mapped monomer encoding* m_i.

Polymer Grid Gate. Inspired by rectangular grid gates, for each polymer $P \in \mathcal{P}(M)$ that can be formed in $\mathcal{A}$, we include $|M|$-many *polymer grid gates* (or simply, *polymer gates*) in $\mathcal{B}$. Although these gates share the same overall structure, the monomers in different polymer gates corresponding to the same polymer either share only a subset of domains or are entirely disjoint; this distinction is described below. An example of a polymer gate for $P = \{m_1, m_1, m_2\}$ is shown in Figure 4, where $|P| = 3$. Each polymer gate consists of three groups of monomers:

1. Polymer Grid Base. The polymer grid base, or *polymer base* for short, is the monomer that serves as the mechanism that forms polymers in the TBN corresponding to target polymers in the TAN. For a TAN polymer P, a polymer base G_P, included in one of the $|M|$-many polymer gates, is constructed as follows. For each monomer $m_i \in P$, we include k copies of the domain m_i^* so that the mapped monomer encoding m_i can bind to the base; we refer to this portion as the *mapped part*. In addition, the polymer base includes the grid base of a rectangular grid gate of width $k|P|$ and height $(k-1)|P| - (u\mathcal{F}(P) + |P| + v)$. The domains used in this grid base are disjoint from any m_i^* domains in $\mathcal{B}$. We denote the height of the polymer base, i.e., the height of the underlying rectangular grid base, by $h(G_P)$. As illustrated in Figure 4(b), the mapped part (forming the first row) and the rectangular grid base (extending from the red to the yellow rows) can be arranged so that the polymer base forms a single rectangular structure. Unlike the mapped part, which may share domains with other polymer bases corresponding to TAN polymers that share monomers with P, the domains in the grid base are unique to G_P. Note that even two polymer bases encoding the same polymer have disjoint sets of domains in their rectangular grid bases. When necessary, we use G_P^j to denote the j-th polymer base where $j \leq |M|$.

2. Base-Specific Horizontals. Each polymer gate includes a collection of horizontal monomers of length $k|P|$ that can bind to the rows of the rectangular grid base in G_P. Since the domains of this grid base are specific to the polymer base, these horizontal monomers are likewise base-specific, and we refer to them as *base-specific horizontals*. Figure 4(c) illustrates the base-specific horizontal monomers for a polymer gate.

3. Base-Specific Verticals. The final group consists of vertical monomers that bind to the columns of the rectangular representation of G_P. More precisely, each vertical monomer contains domains complementary to those in a single column of the grid base, along with exactly one additional domain that binds to the mapped part. For each TAN monomer m_i encoded in the polymer base, there are exactly k vertical monomers containing the domain m_i; in total, there are $k|P|$ base-specific vertical monomers. Figure 4(a) illustrates these base-specific vertical monomers. Note that each vertical monomer contains exactly one domain, its top domain in Figure 4 (a), that binds to the mapped part of the polymer base.

Note that $\mathcal{B}$ is star-limiting. Moreover, since each monomer in $\mathcal{B}$ consists either entirely of starred domains (a *star monomer*) or entirely of unstarred domains (an *unstar monomer*), it follows that in a self-saturated polymer, every starred domain of a star monomer must form a bond with an unstarred domain of an unstar monomer. As each saturated configuration consists of self-saturated polymers, a star monomer (necessarily a polymer base G_P^j) cannot be free in such a configuration. Consequently, the entropy of a saturated configuration is upper bounded by the total number of unstar monomers in the network and a polymer containing d unstar monomers decreases this upper bound by $d-1$. If a polymer includes d unstar monomers, we say that the polymer does $d-1$ unstar merge operations. Then, the entropy of a saturated configuration is equal to the value obtained by subtracting the total number of unstar merge operations across all polymers in the configuration from the total number of unstar monomers in the network.

We will consider two specific polymers: 1) $D_P^{j,1}$ is the polymer that includes the j-th polymer base encoding TAN polymer P, G_P^j, all base-specific horizontal monomers, and, for each $m_i \in P$, the mapped monomer encoding m_i with multiplicity equal to that of m_i in P. 2) $D_P^{j,2}$ is the polymer that includes G_P^j and all base-specific vertical monomers for the polymer base.

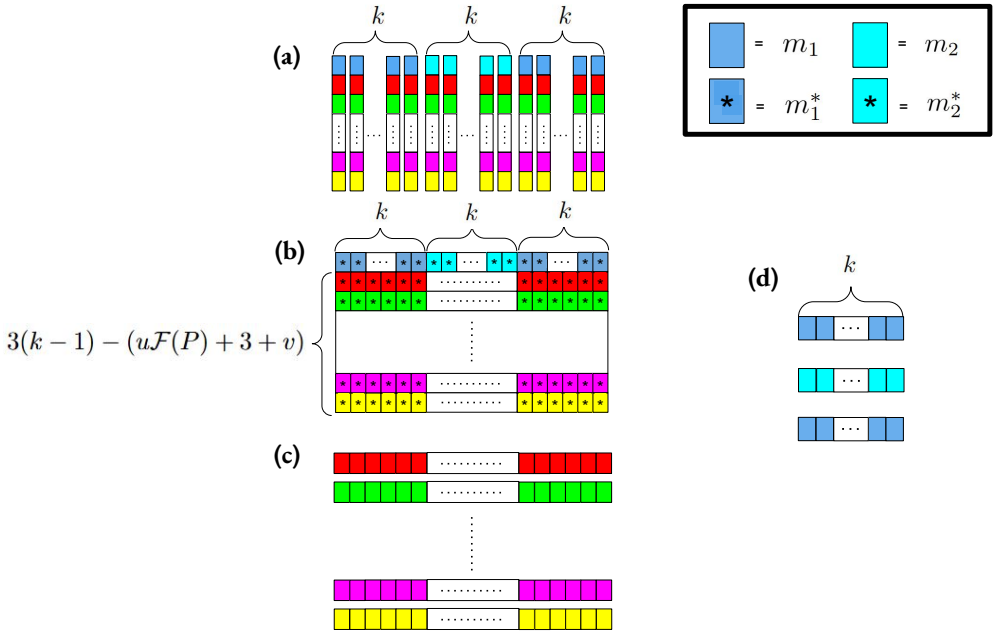


Figure 4 One polymer gate for polymer $\{m_1, m_1, m_2\}$. (a) Base-specific vertical monomers (b) Polymer base encoding the polymer: the first row corresponds to the mapped part, and the remaining rows form the rectangular grid base. (c) Base-specific horizontal monomers (d) Mapped monomers that can bind to the mapped part of the polymer base. The legend inside the black frame shows the color encoding of the domains of mapped monomers and mapped part of the polymer base.

First, note that since neither $D_P^{j,1}$ nor $D_P^{j,2}$ has any exposed starred or unstarred domains, if either is contained in a self-saturated polymer, it can be split off without harming saturation. Furthermore, both polymers include only one star monomer, the polymer base G_P^j. Therefore, if either of these two polymers were splittable, as they have no exposed starred or unstarred domains, after splitting, the resulting polymer containing the polymer base would necessarily have an exposed starred domain and hence would not be self-saturated. Hence, both $D_P^{j,1}$ and $D_P^{j,2}$ are unsplittable polymers. Further note that the polymer $D_P^{j,1}$ includes $h(G_P)+|P| = k|P|-(u\mathcal{F}(P)+|P|+v)$ unstar monomers whereas $D_P^{j,2}$ includes $k|P|$ unstar monomers.

Our first observation, which is characterized by the following lemma, formalizes the local grid-gate-type behavior of each polymer gate in an arbitrary saturated configuration of $\mathcal{B}$:

▶ **Lemma 4.** *In a saturated configuration β of $\mathcal{B}$, if a polymer including a polymer base G_P^j does not contain all the base-specific horizontal monomers, it is either $D_P^{j,2}$ or splittable.*

Proof. If β is a saturated configuration, then every polymer in β is self-saturated; that is, no polymer contains an exposed starred domain. Consequently, in any self-saturated polymer in $\mathcal{B}$, all domains of G_P^j must be bound to an unstarred domain. Now suppose a polymer does not include all base-specific horizontal monomers. Since the domains of the grid base contained in G_P^j are not shared with any other polymer base, it follows by Lemma 3 that the polymer must include all base-specific vertical monomers in order to remain self-saturated. However, this implies that the polymer is either $D_P^{j,2}$ or $D_P^{j,2}$ can be split off without harming saturation. ◀

Next, we will show that no stable configuration of $\mathcal{B}$ can have a polymer including multiple polymer bases within.

▶ **Definition 5.** *We say a polymer in $\mathcal{B}$ is* complete *with respect to a polymer base G_P^j if the polymer includes G_P^j and, for each $m_i \in P$, at least as many mapped monomers encoding m_i as the multiplicity of m_i in P.*

▶ **Lemma 6.** *If there exists a self-saturated polymer P' of $\mathcal{B}$ including multiple polymer bases that is complete with respect to at least one of them, the polymer is splittable and hence cannot appear in a stable configuration.*

Proof. Suppose there exists a stable configuration β of $\mathcal{B}$ including a polymer P' which is complete for G_P^j. By Lemma 4, P' necessarily includes all the base-specific horizontals for G_P^j, since otherwise $S(\beta)$ could be increased by splitting off $D_P^{j,2}$, contradicting β being stable. Then, if P' is complete and includes all base-specific horizontals for G_P^j, $D_P^{j,1}$ can be split off from P' without harming saturation, which again contradicts β being stable. ◀

Lemma 4 and Lemma 6 imply that $\mathcal{B}$ has four types of unsplittable polymers:

1. A single polymer base G_P^j together with all base-specific vertical monomers (necessarily of the form $D_P^{j,2}$, since otherwise $D_P^{j,2}$ can be split off from the polymer).
2. A single polymer base G_P^j together with all base-specific horizontal monomers and complete with respect to G_P^j (necessarily of the form $D_P^{j,1}$, since otherwise $D_P^{j,1}$ can be split off from the polymer).
3. One or more polymer bases, each with all base-specific horizontal monomers, but not complete with respect to any base.
4. A free (unstarred) monomer, possibly a mapped monomer.

Next, we will show that the only way for a polymer of the form (3) to be unsplittable is to include exactly one polymer base:

▶ **Lemma 7.** *If a self-saturated polymer of the form (3) includes multiple polymer bases, the polymer is splittable.*

Proof. For a self-saturated polymer P of the form (3) containing multiple polymer bases, let R denote the multiset of TAN monomers whose mapped monomers are included in P. We construct a polymer P'' that can be split from P, such that both P'' and the remaining polymer P_r are self-saturated.

The polymer P'' consists of the polymer base $G_{P'}^i$, which is contained in P, together with the mapped monomers encoding the monomers in the multiset intersection $P' \cap R$, base-specific horizontal monomers and vertical monomers, base-specific or otherwise, whose unstarred domains correspond to the monomers in $P' \setminus R$. The number of such vertical monomers in P'' is exactly $k|P' \setminus R|$, so that they bind precisely to the starred domains in the mapped portion of $G_{P'}^i$ for which no mapped monomer in P is available to bind. Hence, every starred domain in P'' is saturated, and therefore P'' is self-saturated.

Now consider the remaining polymer P_r. By construction, P_r necessarily contains all base-specific horizontal monomers associated with each polymer base it includes. Consequently, every starred domain in the rectangular grid base of each polymer base can form a bond. Suppose, for contradiction, that P_r is not self-saturated. Then there exists an exposed starred domain in the mapped portion of some polymer base contained in P_r. However, in P'', the number of unstarred domains capable of binding to the mapped portion is exactly equal to the number of starred domains in that mapped portion. Therefore, the existence of such an exposed starred domain in P_r would imply the existence of an exposed starred domain in P, contradicting the assumption that P is self-saturated. Thus, P_r is also self-saturated. ◀

Next, we show that the unsplittable polymers of the form (3) cannot appear in a stable configuration, by exhibiting an incentive to form $D_P^{j,2}$ when some or all of the mapped monomers encoding the monomers in P are absent:

▶ **Lemma 8.** *A stable configuration of TBN $\mathcal{B}$ cannot have a polymer of the form (3).*

Proof. Suppose there exists a stable configuration β of $\mathcal{B}$ such that β includes a polymer of the form (3). We will consider another saturated configuration β'. The polymers of the form (1) and form (2) in β are preserved in β'. On the other hand, β' includes $D_P^{j,2}$ for each G_P^j that is in a polymer of the form (3) in β. Note that the polymers of the form (1) in β only include base-specific vertical monomers for the polymer base they contain, and the polymers of the form (2) do not include any base-specific vertical monomers. Hence, β' can include polymers of the form $D_P^{j,2}$ for each G_P^j that are in a polymer of the form (3) in β while keeping form (1) and form (2) polymers of the same configuration intact.

Take a polymer of the form (3) from β, which, by Lemma 7, includes only a single polymer base G_P^j. As the polymer is self-saturated, each starred domain of a polymer base necessarily forms a bond. Moreover, any monomer other than a mapped monomer that can bind to the mapped part of G_P^j, precisely the vertical monomers, can only form one bond with the mapped part. As the polymer is incomplete for G_P^j, this implies that there are at least k-many additional vertical monomers included in the polymer. Furthermore, any polymer of the form (3) includes all base-specific horizontal monomers. Hence, the total number of unstar monomers in the polymer is at least $h(G_P^j) + k$, for which $h(G_P^j) + k = (k-1)|P| - (u\mathcal{F}(P) + |P| + v) + k = k|P| + k - (u\mathcal{F}(P) + 2|P| + v) > k|P|$ as $k = u \max_P \mathcal{F}(P) + 2|M| + v + 1 > (u\mathcal{F}(P) + 2|P| + v)$. Then a polymer of the form (3) including G_P^j causes a drop in entropy strictly greater than that caused by $D_P^{j,2}$. Hence, $S(\beta) < S(\beta')$ and β cannot be a stable configuration of $\mathcal{B}$. ◀

Lemma 8 implies that the polymers in a stable configuration of $\mathcal{B}$ are of the form (1), (2), or (4). Next, we show that there are further restrictions on the polymers of the form (4).

▶ **Lemma 9.** *No mapped monomer can be free in a stable configuration of $\mathcal{B}$.*

Proof. Suppose there exists a stable configuration β of $\mathcal{B}$ that includes at least one free mapped monomer, say m_1. As $\mathcal{B}$ includes $|M|$-many polymer bases for each TAN polymer, at least one polymer in β that includes a polymer base encoding $P = \{m_1\}$ is necessarily incomplete for the base. Hence, Lemma 8 implies that the polymer is $D_P^{j,2}$. The polymer $D_P^{j,2}$ includes $k|P| = k$-many unstar monomers.

We consider another configuration, β', where the only change from β is that the polymer base under consideration is in the form $D_P^{j,1}$, using the mapped monomer encoding m_1 (which is free in β) and the base-specific horizontal monomers (which are also free in β). The polymer $D_P^{j,1}$ consists of only $h(G_P^j) + |P| = h(G_P^j) + 1$ unstar monomers, where $k - (h(G_P^j) + 1) = u\mathcal{F}(P) + 1 + v$ by the definition of $h(G_P^j)$. Hence, forming $D_P^{j,1}$ uses $u\mathcal{F}(P) + 1 + v$ fewer unstar monomers. As $u, v \in \mathbb{Z}_{\geq 0}$, the term $u\mathcal{F}(P) + 1 + v$ is positive. Hence, $S(\beta') > S(\beta)$, which implies that β cannot be a stable configuration. ◀

Lemma 8 and Lemma 9 imply that in a stable configuration of $\mathcal{B}$, a polymer can be $D_P^{j,1}$, $D_P^{j,2}$, or a free unstar monomer that is not a mapped monomer. Thus, a mapped monomer encoding m_i is in the polymer $D_P^{j,1}$, where $m_i \in P$.

For an arbitrary configuration $\alpha = \{P_1, P_2, ..., P_n\}$ of $\mathcal{A}$, there exists a saturated configuration of $\mathcal{B}$, stable or not, which consists of the polymers $D_{P_1}^{j_1,1}, D_{P_2}^{j_2,1}, ..., D_{P_n}^{j_n,1}$ and $D_P^{j,2}$ for the remaining polymer bases, i.e., for $(j, P) \neq (j_i, P_i)$ for any $i \leq n$. We call such a

configuration a *base configuration for* α, since each mapped monomer is in a polymer of the form $D_P^{j,1}$ that includes a polymer base encoding P. If β is a base configuration for α, $R_{\text{conf}}(\beta) = \alpha$. Moreover, Lemma 4 through Lemma 9 imply that a stable configuration of $\mathcal{B}$ has the same form.

We denote by C the entropy of the configuration of $\mathcal{B}$ in which each polymer base forms the polymer $D_P^{j,2}$. As each formation of a $D_P^{j,1}$ polymer increases C by $u\mathcal{F}(P) + |P| + v$ and a stable configuration maximizes the entropy, the entropy of a stable configuration of $\mathcal{B}$ is equal to $\max_\alpha(C + \sum_{P\in\alpha}(u\mathcal{F}(P) + |P| + v))$.

▶ **Lemma 10.** *For each stable configuration β of $\mathcal{B}$, $R_{conf}(\beta)$ is a stable configuration of $\mathcal{A}$.*

Proof. Let β be a stable configuration of $\mathcal{B}$. As $\sum_{P\in\alpha}|P| = |M|$ for an arbitrary configuration α of $\mathcal{A}$, the expression for the entropy of a stable configuration of $\mathcal{B}$ can be rewritten as $C + |M| + \max_\alpha(\sum_{P\in\alpha}(u\mathcal{F}(P) + v))$. As β is a stable configuration, $R_{\text{conf}}(\beta)$ maximizes $\sum_{P\in\alpha}(u\mathcal{F}(P) + v)$. Hence, $R_{\text{conf}}(\beta) = \arg\max_\alpha \sum_{P\in\alpha}(u\mathcal{F}(P) + v) = \arg\max_\alpha(u(\sum_{P\in\alpha}\mathcal{F}(P)) + v|\alpha|) = \arg\max_\alpha(uH(\alpha) + vS(\alpha)) = \arg\max_\alpha(-E_{\text{TAN}}(\alpha)) = \arg\min_\alpha E_{\text{TAN}}(\alpha)$. Hence, $R_{\text{conf}}(\beta)$ is a stable configuration of $\mathcal{A}$. ◀

▶ **Lemma 11** (Relative energies)**.** *Given TAN $\mathcal{A} = (M, \mathcal{F}, u, v)$, let R_{conf} be the mapping from the TBN $\mathcal{B}$ to $\mathcal{A}$, constructed as above, $E_\mathcal{A}$ be the energy of a stable configuration of $\mathcal{A}$ and $S_\mathcal{B}$ be the entropy of a stable configuration of $\mathcal{B}$. For every configuration α of $\mathcal{A}$, there exists a saturated configuration $\beta \in R_{conf}^{-1}(\alpha)$ of $\mathcal{B}$ such that $S_\mathcal{B} - S(\beta) = E_{TAN}(\alpha) - E_\mathcal{A}$.*

Proof. For a configuration $\alpha = \{P_1, P_2, ..., P_n\}$ of $\mathcal{A}$, consider β which is a base configuration for α. Note that $R_{\text{conf}}(\beta) = \alpha$. Therefore, $S(\beta) = C + |M| + \sum_{P\in\alpha}(u\mathcal{F}(P) + v)$. Furthermore, $S_\mathcal{B} = C + |M| + \max_{\alpha'}\sum_{P\in\alpha'}(u\mathcal{F}(P) + v)$. As $E_\mathcal{A} = \min_{\alpha'} -\sum_{P\in\alpha'}(u\mathcal{F}(P) + v)$ and $E(\alpha) = -\sum_{P\in\alpha}(u\mathcal{F}(P) + v)$, it follows that $S_\mathcal{B} - S(\beta) = \max_{\alpha'}\sum_{P\in\alpha'}(u\mathcal{F}(P) + v) - \sum_{P\in\alpha}(u\mathcal{F}(P) + v) = -E_\mathcal{A} + E(\alpha) = E_{\text{TAN}}(\alpha) - E_\mathcal{A}$. ◀

As any stable configuration α of $\mathcal{A}$ satisfies $E_\mathcal{A} = E(\alpha)$, the immediate corollary of Lemma 11 is the following:

▶ **Corollary 12.** *For every stable configuration α of $\mathcal{A}$, there exists a stable configuration β of $\mathcal{B}$ where $\beta \in R_{conf}^{-1}(\alpha)$.*

▶ **Theorem 13.** *For every TAN $\mathcal{A}$ with the simplified energy function, there exists a TBN $\mathcal{B}$ such that $\mathcal{B}$ indirectly realizes $\mathcal{A}$.*

Proof. The result is a direct consequence of Lemma 10 and Corollary 12. ◀

6 Realizing TANs with TBNs: Extending to the Full Energy Model

We now present the construction of the realizing TBN for the general TAN energy function:

$$E_{\text{TAN}}(\alpha) = -uH(\alpha) - vS(\alpha) + \sum_{P_s\in\text{supp}(\alpha)}\sum_{j=1}^{\underline{\alpha}(s)}\log j$$

We assume that each value of $\log j$, for $j \in \{1, 2, \ldots, |M|\}$, is represented with 5-digit precision and hence is rational. Since $u, v \in \mathbb{Q}_{\geq 0}$ and $\mathcal{F}$ is a non-negative rational valued function, there exists a smallest $C_{\text{TAN}} \in \mathbb{N}$ such that $C_{\text{TAN}}(u\mathcal{F}(P) + v + (\log|M| + 1)|P| - \log j) \in \mathbb{N}$ for every $P \in \mathcal{P}(M)$ and $j \in \{1, 2, \ldots, |M|\}$.

In contrast to the construction of Section 5, the new construction again consists of polymer gates and mapped monomers, and for every polymer P of the realized TAN $\mathcal{A}$, there are $|M|$ polymer bases G_P^j. However, the height of each polymer base now depends on its index. Specifically, for each $j \in \{1, 2, \ldots, |M|\}$, we define $h(G_P^j) = (k-1)|P| - C_{\text{TAN}}(u\mathcal{F}(P) + v + (\log|M|+1)|P| - \log j)$. Thus, unlike the construction in Section 5, the polymer bases associated with a fixed polymer P are no longer uniform in height. We further define $k = |M| + C_{\text{TAN}}(u \max_{P\in\mathcal{P}(M)} \mathcal{F}(P) + v + (\log|M|+1)|M|) + 1$. Note that the definitions of C_{TAN} and k imply that $h(G_P^j) \in \mathbb{N}$.

The definition of $h(G_P^j)$ implies that the number of unstarred monomers in $D_P^{j,1}$, the polymer in which G_P^j is saturated by base-specific horizontal monomers and mapped monomers, is $h(G_P^j) + |P| = k|P| - C_{\text{TAN}}(u\mathcal{F}(P) + v + (\log|M|+1)|P| - \log j)$. Thus, compared to $D_P^{j,2}$, the polymer in which G_P^j is saturated by base-specific vertical monomers, the polymer $D_P^{j,1}$ contains $C_{\text{TAN}}(u\mathcal{F}(P) + v + (\log|M|+1)|P| - \log j)$ fewer unstarred monomers.

The choice of the height function $h(G_P^j)$ is motivated by the behavior of the combinatorial entropy term when polymers $D_P^{j,1}$ are formed for consecutive indices beginning with $j = 1$. In particular, the decrease arising from the entropy contribution of multiple copies of the same polymer is captured by the sum of the logarithmic terms appearing in the definition of $h(G_P^j)$. In Corollary 17, we show that this behavior is realized in the stable configurations of the construction.

Note that the new construction differs from the construction of Section 5 only in the definitions of $h(G_P^j)$ and k, while the overall structure of the construction remains unchanged. Consequently, Lemma 4 and Lemma 6 continue to hold immediately for the new construction, reducing the possible unsplittable polymers of the network to those of forms (1), (2), (3), and (4). Moreover, as in the original construction, unsplittable polymers of form (3) are further restricted by Lemma 7.

Accordingly, we focus on the new results, analogous to those from Lemma 8 through Theorem 13, which establish the intended behavior in the presence of the additional terms arising from combinatorial entropy. Throughout this section, we again denote the realized TAN by $\mathcal{A}$ and the realizing TBN by $\mathcal{B}$.

▶ **Lemma 14.** *A stable configuration of $\mathcal{B}$ cannot include a polymer of the form (3).*

Proof. Let G_P^j be the unique polymer base contained in an unsplittable polymer of form (3). Since the polymer is incomplete for P and contains all base-specific horizontal monomers, the number of unstarred monomers contained in the polymer is at least $h(G_P^j) + k = k|P| - |P| - C_{\text{TAN}}(u\mathcal{F}(P) + v + (\log|M|+1)|P| - \log j) + k$. Because $\log j \geq 0$ for all $j \geq 1$, the definition of k implies that $k - |P| - C_{\text{TAN}}(u\mathcal{F}(P) + v + (\log|M|+1)|P| - \log j) > 0$. Therefore, $h(G_P^j) + k > k|P|$, showing that a polymer of form (3) contains strictly more unstarred monomers than the polymer $D_P^{j,2}$.

The remainder of the argument proceeds analogously to Lemma 8. Suppose a configuration β of $\mathcal{B}$ contains a polymer of form (3). Then one can construct another configuration β' by replacing every such polymer with the corresponding polymer $D_P^{j,2}$, in which the polymer base is saturated using base-specific vertical monomers. This replacement strictly increases the number of polymers in the configuration, implying that $S(\beta) < S(\beta')$. This contradicts the assumption that β is stable. ◀

Lemma 14 implies that the only polymers that can be included in a stable configuration of $\mathcal{B}$ are of the form (1), (2), or (4). The result analogous to Lemma 9 is the following:

▶ **Lemma 15.** *A stable configuration of $\mathcal{B}$ cannot include a free mapped monomer.*

Proof. Similar to the proof of Lemma 9, suppose there exists a configuration β of $\mathcal{B}$ that contains a free mapped monomer m_1. Then there exists a polymer base G^j_P, where $P = \{m_1\}$, that is included in a polymer of the form $D^{j,2}_P$.

Now consider a configuration β' obtained from β by keeping all polymers unchanged except that $D^{j,2}_P$ is replaced by $D^{j,1}_P$. Since $h(G^j_P)+|P| = k|P|-C_{\text{TAN}}(u\mathcal{F}(P)+v+(\log|M|+1)|P|-\log j)$ and $|P| = 1$, it follows that $S(\beta')-S(\beta) = C_{\text{TAN}}(u\mathcal{F}(P)+v+\log|M|+1-\log j)$, which is strictly positive, since $j \in \{1, 2, \ldots, |M|\}$. Moreover, since $\mathcal{F}(P)$, u, and v are all non-negative, it follows that $S(\beta') > S(\beta)$, contradicting the stability of β. ◀

Next, we will show that the desired accumulating effect on the logarithmic terms is a necessary condition at stability:

▶ **Lemma 16.** *A stable configuration of $\mathcal{B}$ cannot include both $D^{j,2}_P$ and $D^{j+1,1}_P$ for $j \in \{1, 2, ..., |M| - 1\}$.*

Proof. Suppose there exists a stable configuration β of $\mathcal{B}$ that contains both $D^{j,2}_P$ and $D^{j+1,1}_P$. The total number of unstarred monomers contained in these two polymers is $2k|P| - C_{\text{TAN}}(u\mathcal{F}(P) + v + (\log|M| + 1)|P| - \log(j + 1))$.

We construct another configuration β', obtained from β by replacing the polymer bases under consideration so that they form polymers of type $D^{j,1}_P$ and $D^{j+1,2}_P$, respectively, instead of $D^{j,2}_P$ and $D^{j+1,1}_P$. The total number of unstarred monomers in $D^{j,1}_P$ and $D^{j+1,2}_P$ is then $2k|P| - C_{\text{TAN}}(u\mathcal{F}(P) + v + (\log|M| + 1)|P| - \log j)$.

Hence, $S(\beta') - S(\beta) = C_{\text{TAN}}(\log(j + 1) - \log j)$, which is strictly positive. Therefore, β cannot be a stable configuration. ◀

As Lemma 14 implies that each polymer base is necessarily either in the polymer $D^{j,1}_P$ or $D^{j,2}_P$ at stability, the immediate corollary of Lemma 16 is the following:

▶ **Corollary 17.** *In a stable configuration of $\mathcal{B}$, for each polymer P of $\mathcal{A}$, one of the following is true:*

- *G^j_P is in the polymer $D^{j,2}_P$ for all $j \in \{1, 2, ..., |M|\}$*
- *There exists $j' \in \{1, 2, ..., |M|\}$ such that each polymer base G^j_P for $j \leq j'$ is in the polymer $D^{j,1}_P$ and each polymer base G^j_P for $j' < j$ is in the polymer $D^{j,2}_P$.*

For an arbitrary configuration α of $\mathcal{A}$, where $\text{supp}(\alpha) = \{P_1, P_2, \ldots, P_n\}$, we define the saturated configuration of $\mathcal{B}$, stable or not, which, for each $s \in \{1, 2, \ldots, n\}$, consists of the polymers $D^{j,1}_{P_s}$ for $j \in \{1, 2, \ldots, \underline{\alpha}(s)\}$ and $D^{j,2}_{P_s}$ for $j > \underline{\alpha}(s)$, as the *base configuration for* α (all polymer bases encoding polymers that are not in $\text{supp}(\alpha)$ form $D^{j,2}_P$). In contrast to the construction of Section 5, the base configuration for α is unique, since the present construction explicitly depends on the indices of the polymer bases. Furthermore, if β is the base configuration for α, then $R_{\text{conf}}(\beta) = \alpha$. Hence, R_{conf} defines a bijection from the set of base configurations of $\mathcal{B}$ to the set of all configurations of $\mathcal{A}$.

Moreover, the fact that the base configuration contains polymers $D^{j,1}_{P_s}$ for consecutive indices beginning at $j = 1$ and ending at $j = \underline{\alpha}(s)$ is consistent with Corollary 17. Together, Lemma 14 through Corollary 17 imply that every stable configuration of $\mathcal{B}$ is necessarily a base configuration.

As in Section 5, we denote by C_2 the entropy of the configuration of $\mathcal{B}$ in which every polymer base forms the polymer $D^{j,2}_P$. Each formation of a polymer $D^{j,1}_P$ increases this entropy by $C_{\text{TAN}}(u\mathcal{F}(P) + v + (\log|M| + 1)|P| - \log j)$. Since the base configuration for α contains polymers $D^{j,1}_P$ for consecutive indices, the logarithmic terms accumulate. Therefore, if β is the

base configuration corresponding to α, then $S(\beta) = C_2 + C_{\text{TAN}} \sum_{P_s \in \text{supp}(\alpha)} (\underline{\alpha}(s) u \mathcal{F}(P_s) + \underline{\alpha}(s) v + \underline{\alpha}(s)(\log|M|+1)|P_s| - \sum_{j=1}^{\underline{\alpha}(s)} \log j)$. Using the identity $\sum_{P_s \in \text{supp}(\alpha)} \underline{\alpha}(s)|P_s| = |M|$, it follows that $S(\beta) = C_2 + C_{\text{TAN}}(\log|M|+1)|M| - C_{\text{TAN}} E_{\text{TAN}}(\alpha)$.

The fact that every stable configuration of $\mathcal{B}$ is necessarily a base configuration, together with the fact that R_{conf} is a bijection when restricted to the set of base configurations of $\mathcal{B}$, allows us to establish both directions of the thermodynamic realization of $\mathcal{A}$ by $\mathcal{B}$.

▶ **Lemma 18.** *For each stable configuration β of $\mathcal{B}$, $R_{conf}(\beta)$ is a stable configuration of $\mathcal{A}$.*

Proof. Suppose that β is a stable configuration of $\mathcal{B}$ but $R_{\text{conf}}(\beta) = \alpha$ is not a stable configuration of $\mathcal{A}$. Then there exists α' such that $E_{\text{TAN}}(\alpha') < E_{\text{TAN}}(\alpha)$. Let β' be the base configuration for α'. Then $S(\beta') = C_2 + C_{\text{TAN}}(\log|M|+1)|M| - C_{\text{TAN}} E_{\text{TAN}}(\alpha') > C_2 + C_{\text{TAN}}(\log|M|+1)|M| - C_{\text{TAN}} E_{\text{TAN}}(\alpha) = S(\beta)$, which contradicts β being a stable configuration of $\mathcal{B}$. ◀

▶ **Lemma 19.** *For every stable configuration α of $\mathcal{A}$, there exists a stable configuration β of $\mathcal{B}$ where $\beta \in R_{conf}^{-1}(\alpha)$.*

Proof. Suppose α is a stable configuration of $\mathcal{A}$ but β, the base configuration for α is not a stable configuration of $\mathcal{B}$. Then there exists a base configuration β', where $R_{\text{conf}}(\beta') = \alpha'$, such that $S(\beta') = C_2 + C_{\text{TAN}}(\log|M|+1)|M| - C_{\text{TAN}} E_{\text{TAN}}(\alpha') > C_2 + C_{\text{TAN}}(\log|M|+1)|M| - C_{\text{TAN}} E_{\text{TAN}}(\alpha)$, which implies $E_{\text{TAN}}(\alpha') < E_{\text{TAN}}(\alpha)$, which contradicts α being a stable configuration of $\mathcal{A}$. ◀

▶ **Theorem 20.** *For every TAN $\mathcal{A}$, there exists a TBN $\mathcal{B}$ such that $\mathcal{B}$ indirectly realizes $\mathcal{A}$.*

Proof. The result is a direct consequence of Lemma 18 and Lemma 19. ◀

Next, we show that a set of TBN monomers obtained via our construction can, without modifying the monomer types, indirectly realize a family of TANs that share the same underlying monomer set but differ in their monomer counts. In other words, the realizing TBN is fixed at the level of monomer types, and only the multiplicities are varied to realize different target TANs.

Before presenting the result, we define a *TAN template* as a 4-tuple $(M_{\text{mon}}^{\text{TAN}}, \mathcal{F}, u, v)$, where $M_{\text{mon}}^{\text{TAN}}$ denotes the set of monomer types. A TAN template specifies a family of TANs obtained by assigning different count vectors to the fixed monomer set. That is, each member TAN is determined by a count vector $\vec{m}_{\text{count}}^{\text{TAN}}$, which specifies the number of copies of each monomer type in $M_{\text{mon}}^{\text{TAN}}$, or equivalently $\vec{m}_{\text{count}}^{\text{TAN}} \in \mathbb{N}^{|M_{\text{mon}}^{\text{TAN}}|}$. Note that, for a fixed TAN template, the pair $(M_{\text{mon}}^{\text{TAN}}, \vec{m}_{\text{count}}^{\text{TAN}})$ uniquely determines a TAN instance. We use an analogous notation for TBNs, writing the underlying monomer set of a TBN as $M_{\text{mon}}^{\text{TBN}}$, and representing a particular TBN instance by the pair $(M_{\text{mon}}^{\text{TBN}}, \vec{m}_{\text{count}}^{\text{TBN}})$. We denote the case in which each entry of $\vec{m}_{\text{count}}^{\text{TAN}}$ is bounded above by a fixed value M' as $\vec{m}_{\text{count}}^{\text{TAN}} \leq M'$.

▶ **Theorem 21.** *For each TAN template $(M_{\text{mon}}^{\text{TAN}}, \mathcal{F}, u, v)$, there exists a set of TBN monomers $M_{\text{mon}}^{\text{TBN}}$ such that for every count vector $\vec{m}_{\text{count}}^{\text{TAN}}$ satisfying $\vec{m}_{\text{count}}^{\text{TAN}} \leq M'$, there exist TBN monomer counts $\vec{m}_{\text{count}}^{\text{TBN}}$ such that the TBN $(M_{\text{mon}}^{\text{TBN}}, \vec{m}_{\text{count}}^{\text{TBN}})$ indirectly realizes the TAN $(M_{\text{mon}}^{\text{TAN}}, \vec{m}_{\text{count}}^{\text{TAN}})$.*

Proof. For a given TAN template $(M_{\text{mon}}^{\text{TAN}}, \mathcal{F}, u, v)$, where $|M_{\text{mon}}^{\text{TAN}}| = n$, let P' be a polymer maximizing $\mathcal{F}$ among all polymers over $M_{\text{mon}}^{\text{TAN}}$ with at most nM' monomers.

We construct the set of TBN monomers $M_{\text{mon}}^{\text{TBN}}$ as the collection of indexed polymer gates (i.e., polymer bases, base-specific horizontal monomers, and base-specific vertical monomers) for all polymers over $M_{\text{mon}}^{\text{TAN}}$ containing at most M' copies of each monomer, together with mapped monomers encoding each monomer in $M_{\text{mon}}^{\text{TAN}}$. We define $h(G_P^j) = (k-1)|P| - C_{\text{unif}}(u\mathcal{F}(P) + v + (\log(nM')+1)|P| - \log j)$, and $k = (nM') + C_{\text{unif}}(u\mathcal{F}(P') + v + (\log(nM')+1)nM') + 1$. Let C_{unif} be the smallest natural number such that $C_{\text{unif}}(u\mathcal{F}(P) + v + (\log(nM')+1)|P| - \log j)$ is a natural number for every P with $|P| \leq nM'$ and every $j \in \{1, 2, \ldots, nM'\}$, where each $\log j$ is taken to 5-digit precision.

Given a TAN instance specified by $\vec{m}_{\text{count}}^{\text{TAN}}$, we instantiate the corresponding TBN counts as follows. First, we include mapped monomers encoding each monomer in $M_{\text{mon}}^{\text{TAN}}$ with multiplicities given by $\vec{m}_{\text{count}}^{\text{TAN}}$. Second, for each polymer over $M_{\text{mon}}^{\text{TAN}}$ compatible with these counts, we include $\sum_i \vec{m}_{\text{count}}^{\text{TAN}}(i)$ copies of the corresponding polymer gates, with heights determined by the above definition of $h(G_P^j)$.

Note that the resulting TBN is structurally identical to the construction used for realizing a fixed-count TAN, except for the specific values of $h(G_P^j)$ and k. These choices are uniform over all TAN instances with counts of each monomer bounded by M', and correctness follows by adapting the proofs of Lemma 14 and Lemma 15. In particular, since P' maximizes $\mathcal{F}$ over all polymers of size at most nM', we obtain $h(G_P^j) + k > k|P|$ and thus the analogue of Lemma 14 holds. Moreover, since $j \leq \sum_i \vec{m}_{\text{count}}^{\text{TAN}}(i) \leq nM'$, we have $\log j \leq \log(nM')$, from which the analogue of Lemma 15 follows. Finally, the arguments in Lemma 16 through Theorem 20 do not depend on monomer counts and therefore carry over directly. Hence, the constructed TBN thermodynamically realizes the given TAN. ◀

7 Discussion

Initially, TBNs looked too restricted to express arbitrary polymer-level preferences. Their only source of enthalpic favorability is pairwise complementary binding, so attaching a reward to a whole complex (rather than to the bonds within it) seemed out of reach. Our main result shows that this intuition is misleading and that every TAN, however its affinities are assigned, has a TBN whose stable configurations match it. The construction works by encoding the polymer-level rewards indirectly: when a polymer forms, it releases a calibrated number of helper monomers, matching the favorability described by the specific TAN it is realizing. Moreover, the construction has a limited form of uniformity: for any bounded family of TANs (that differ only by counts of monomers), one fixed TBN can realize the family by changing only multiplicities. In this way, TANs are not a more powerful model; rather, they are a higher-level language for the thermodynamically equivalent model of TBNs. An important caveat is that our equivalence means that we match stable, minimum-free-energy configurations, not full kinetics or Boltzmann distributions. A natural next step is to aim for capturing a more complete notion of equilibrium thermodynamics.

Towards that objective, we note that Lemma 11 (and the analogous entropy identity in Section 6) shows that the correspondence between TANs and TBNs is stronger than a mere agreement on which configurations are stable. For a TAN configuration α, the construction produces a saturated TBN configuration β mapping to α such that the gap from stability in the TAN is reflected in the gap from maximal entropy in the TBN. Thus, the construction preserves relative thermodynamic preferences, not just equilibrium outcomes. A direct line of future work is expanding our definitions of system equivalence. A "complete" thermodynamic realization may consider preserving all relative energies of a system, rather

than just stable configurations. Developing stronger notions of realization may also be a useful step toward comparing kinetic behavior as well, possibly allowing us to establish more traditional simulation results.

Lastly, because Thermodynamic Binding Networks were introduced as a domain-level abstraction inspired by DNA binding, one can naturally interpret each TBN domain as a designed DNA domain and each monomer as a multi-domain strand or complex [7, 18]. Under this interpretation, one implication of our equivalence result is that any TAN-style equilibrium polymer preference can, in principle, be compiled into DNA-level binding components whose stable complexes match the intended minimum-free-energy states. This perspective aligns with the broader tradition of DNA universality, where strand-displacement systems have already been shown to be universal for implementing arbitrary chemical kinetics [17]. Accordingly, this paper supports the view that DNA may be regarded as a conceptually complete substrate for equilibrium monomer-polymer thermodynamics, although further investigation is needed.

References

1 Gagan Aggarwal, Qi Cheng, Michael H Goldwasser, Ming-Yang Kao, Pablo Moisset De Espanes, and Robert T Schweller. Complexities for generalized models of self-assembly. *SIAM Journal on Computing*, 34(6):1493–1515, 2005. doi:10.1137/S0097539704445202.

2 Marco Antoniotti, Carla Piazza, Alberto Policriti, Marta Simeoni, and Bud Mishra. Taming the complexity of biochemical models through bisimulation and collapsing: theory and practice. *Theoretical Computer Science*, 325(1):45–67, 2004. doi:10.1016/J.TCS.2004.03.064.

3 Keenan Breik, Cameron Chalk, David Doty, David Haley, and David Soloveichik. Programming substrate-independent kinetic barriers with thermodynamic binding networks. *IEEE/ACM transactions on computational biology and bioinformatics*, 18(1):283–295, 2019.

4 Keenan Breik, Chris Thachuk, Marijn Heule, and David Soloveichik. Computing properties of stable configurations of thermodynamic binding networks. *Theoretical Computer Science*, 785:17–29, 2019. doi:10.1016/J.TCS.2018.10.027.

5 Sven De Vries and Rakesh V Vohra. Combinatorial auctions: A survey. *INFORMS Journal on computing*, 15(3):284–309, 2003. doi:10.1287/IJOC.15.3.284.16077.

6 Robert M Dirks, Justin S Bois, Joseph M Schaeffer, Erik Winfree, and Niles A Pierce. Thermodynamic analysis of interacting nucleic acid strands. *SIAM review*, 49(1):65–88, 2007. doi:10.1137/060651100.

7 David Doty, Trent A Rogers, David Soloveichik, Chris Thachuk, and Damien Woods. Thermodynamic binding networks. In *DNA Computing and Molecular Programming: 23rd International Conference, DNA 23, Austin, TX, USA, September 24–28, 2017, Proceedings 23*, pages 249–266. Springer, 2017. doi:10.1007/978-3-319-66799-7_16.

8 Jean-Claude Fernandez. An implementation of an efficient algorithm for bisimulation equivalence. *Science of Computer Programming*, 13(2-3):219–236, 1990. doi:10.1016/0167-6423(90)90071-K.

9 Bin Fu, Matthew J Patitz, Robert T Schweller, and Robert Sheline. Self-assembly with geometric tiles. In *International Colloquium on Automata, Languages, and Programming*, pages 714–725. Springer, 2012. doi:10.1007/978-3-642-31594-7_60.

10 David Haley and David Doty. Computing properties of thermodynamic binding networks: An integer programming approach. *arXiv preprint arXiv:2011.10677*, 2020. arXiv:2011.10677.

11 Robert Johnson, Qing Dong, and Erik Winfree. Verifying chemical reaction network implementations: a bisimulation approach. *Theoretical Computer Science*, 765:3–46, 2019. doi:10.1016/J.TCS.2018.01.002.

12 Maxim P Nikitin. Non-complementary strand commutation as a fundamental alternative for information processing by dna and gene regulation. *Nature Chemistry*, 15(1):70–82, 2023.

13 Tosan Omabegho. Allosteric linkages that model the chemical cycle of a molecular motor enzyme. *bioRxiv*, pages 2021–04, 2021.

14 Joshua Petrack, David Soloveichik, and David Doty. Thermodynamically driven signal amplification. In *29th International Conference on DNA Computing and Molecular Programming*, page 1, 2023.

15 Talal Rahwan, Tomasz P Michalak, Michael Wooldridge, and Nicholas R Jennings. Coalition structure generation: A survey. *Artificial Intelligence*, 229:139–174, 2015. doi:10.1016/J.ARTINT.2015.08.004.

16 Jason W Rocks, Nidhi Pashine, Irmgard Bischofberger, Carl P Goodrich, Andrea J Liu, and Sidney R Nagel. Designing allostery-inspired response in mechanical networks. *Proceedings of the National Academy of Sciences*, 114(10):2520–2525, 2017.

17 David Soloveichik, Georg Seelig, and Erik Winfree. DNA as a universal substrate for chemical kinetics. *Proceedings of the National Academy of Sciences*, 107(12):5393–5398, 2010.

18 Boya Wang, Cameron Chalk, David Doty, and David Soloveichik. Molecular computation at equilibrium via programmable entropy. *Science Advances*, 12(2):eadx3969, 2026.

19 Erik Winfree. *Algorithmic self-assembly of DNA*. California Institute of Technology, 1998.

A Derivation of the Macrostate Energy Function

In Section 2 we assign to each configuration α the energy $-\log\Omega(\alpha)$, where $\Omega(\alpha)$ is the number of microstates consistent with the configuration α. This appendix derives that expression from first principles, by counting the microstates of an ideal dilute solution. We measure energy in units of k_BT and entropy in units of k_B, so that $\log\Omega$ is an entropy and $-\log\Omega(\alpha)$ a dimensionless free energy; log denotes the natural logarithm. The derivation is based on [6], but is standard in the statistical mechanics literature.

As in Section 2, let M be the finite multiset of *monomer* types and $\mathcal{P}(M)$ the set of *polymer* types, each a nonempty sub-multiset of M. A *macrostate* is a configuration α, that is, a partition of M into polymers; equivalently, it records the molecular count $\underline{\alpha}(s)$ of each polymer type $P_s \in \mathrm{supp}(\alpha) = \{P_1, P_2, \ldots, P_n\}$. Let $|\alpha| = \sum_{P_s\in\mathrm{supp}(\alpha)} \underline{\alpha}(s)$ denote the total number of polymers. We model the solution as V cells, where V is the volume in units of a reference cell, and we work in the dilute regime $V \gg |\alpha|$, so that polymers occupy cells independently (the ideal-solution assumption).

For each polymer type P, let Ω_P count the microstates of the universe that are made available by assembling a single polymer of type P from its constituent free monomers. These microstates include internal microstates of the polymer, such as different conformations, as well as those arising from heat released into the thermal reservoir due to the formation of bonds.

The number $\Omega(\alpha)$ of microstates factors into four contributions:

(A) Positioning the polymers. Treating the $|\alpha|$ polymers as distinguishable objects, each independently occupies one of V cells: $A = V^{|\alpha|}$.

(B) Making the monomers distinguishable. Writing μ_a for the number of copies of monomer type a in M, labeling the monomers of each type so that they may be told apart contributes $B = \prod_a \mu_a!$, the product ranging over monomer types.

(C) Indistinguishability of polymers. Polymers of the same type are interchangeable, so the relabelings of identical polymers must be removed: $C = \dfrac{1}{\prod_{P_s\in\mathrm{supp}(\alpha)} \underline{\alpha}(s)!}$.

(D) Internal microstates. Each polymer of type P_s independently realizes one of Ω_{P_s} microstates: $D = \prod_{P_s\in\mathrm{supp}(\alpha)} \Omega_{P_s}^{\underline{\alpha}(s)}$.

Collecting the four factors,

$$\Omega(\alpha) = A\cdot B\cdot C\cdot D = V^{|\alpha|}\left(\prod_a \mu_a!\right)\frac{1}{\prod_{P_s\in\mathrm{supp}(\alpha)} \underline{\alpha}(s)!}\prod_{P_s\in\mathrm{supp}(\alpha)} \Omega_{P_s}^{\underline{\alpha}(s)}.$$

For the fixed monomer multiset M, all microstates of the universe are equiprobable, so the probability of a configuration α is proportional to $\Omega(\alpha)$. Thus, the most probable configurations are those that minimize $-\log\Omega(\alpha)$, which we therefore adopt as the energy. The factor (B) depends only on the fixed multiset M and not on the configuration α, so it shifts $-\log\Omega(\alpha)$ by a constant and is irrelevant to the minimization. (Equivalently, whether the monomers are deemed distinguishable has no effect on equilibrium, the familiar Gibbs resolution of the mixing paradox.) Dropping this constant leaves

$$\Omega(\alpha) = \frac{V^{|\alpha|}}{\prod_{P_s\in\mathrm{supp}(\alpha)} \underline{\alpha}(s)!} \prod_{P_s\in\mathrm{supp}(\alpha)} \Omega_{P_s}^{\underline{\alpha}(s)}, \qquad \text{up to a configuration-independent factor.}$$

Because the regimes of interest often involve only a few copies of each polymer, we retain the factorial exactly rather than approximating it via Stirling's formula $\log(n!) = n\log n - n + O(\log n)$ (as in [6]), which is accurate only when every count $\underline{\alpha}(s)$ is large. Writing each $\log\big(\underline{\alpha}(s)!\big)$ as a direct sum of logarithms,

$$\log\big(\underline{\alpha}(s)!\big) = \log\underline{\alpha}(s) + \log\big(\underline{\alpha}(s)-1\big) + \cdots + \log 2 = \sum_{j=1}^{\underline{\alpha}(s)} \log j,$$

an empty sum equal to 0 when $\underline{\alpha}(s) \le 1$, then taking logarithms of $\Omega(\alpha)$ and negating gives the exact macrostate energy,

$$-\log\Omega(\alpha) = -\sum_{P_s\in\mathrm{supp}(\alpha)} \underline{\alpha}(s)\log V \; - \sum_{P_s\in\mathrm{supp}(\alpha)} \underline{\alpha}(s)\log\Omega_{P_s} \; + \sum_{P_s\in\mathrm{supp}(\alpha)} \sum_{j=1}^{\underline{\alpha}(s)} \log j, \quad (1)$$

valid for every configuration, whatever the counts. Identifying $\log V = v$ and $\log\Omega_{P_s} = u\,\mathcal{F}(P_s)$ with $u = w\log V$, and recalling that $S(\alpha) = |\alpha|$ is the number of polymers and $H(\alpha) = \sum_{P_s\in\mathrm{supp}(\alpha)} \underline{\alpha}(s)\,\mathcal{F}(P_s)$ the enthalpy, the first two sums become $-vS(\alpha)$ and $-uH(\alpha)$, so that

$$-\log\Omega(\alpha) = -uH(\alpha) - vS(\alpha) + \sum_{P_s\in\mathrm{supp}(\alpha)} \sum_{j=1}^{\underline{\alpha}(s)} \log j = E_{\mathrm{TAN}}(\alpha),$$

the energy of Section 2, whose final term is the combinatorial entropy.