Lattice Path Conference
20-24 July 2026
TU Wien
Vienna, Austria

Lattice paths and the enumeration for stacks of protein contact maps

Lisa H. Sun (Joint work with Q.-H. Guo, Y.L. Jin, C.Y. Li, and S.N. Xu)
Center for Combinatorics, LPMC, Nankai University, Tianjin 300071, P.R. China

Protein contact maps

In the lattice model, two amino acid residues of a protein fold form a contact if they are not consecutive in the chain, and reside on two adjacent points in the lattice. All contacts of a protein fold form the contact map, which can be represented by placing the residues horizontally and drawing an arc in the upper-half plane between two residues which form a contact.

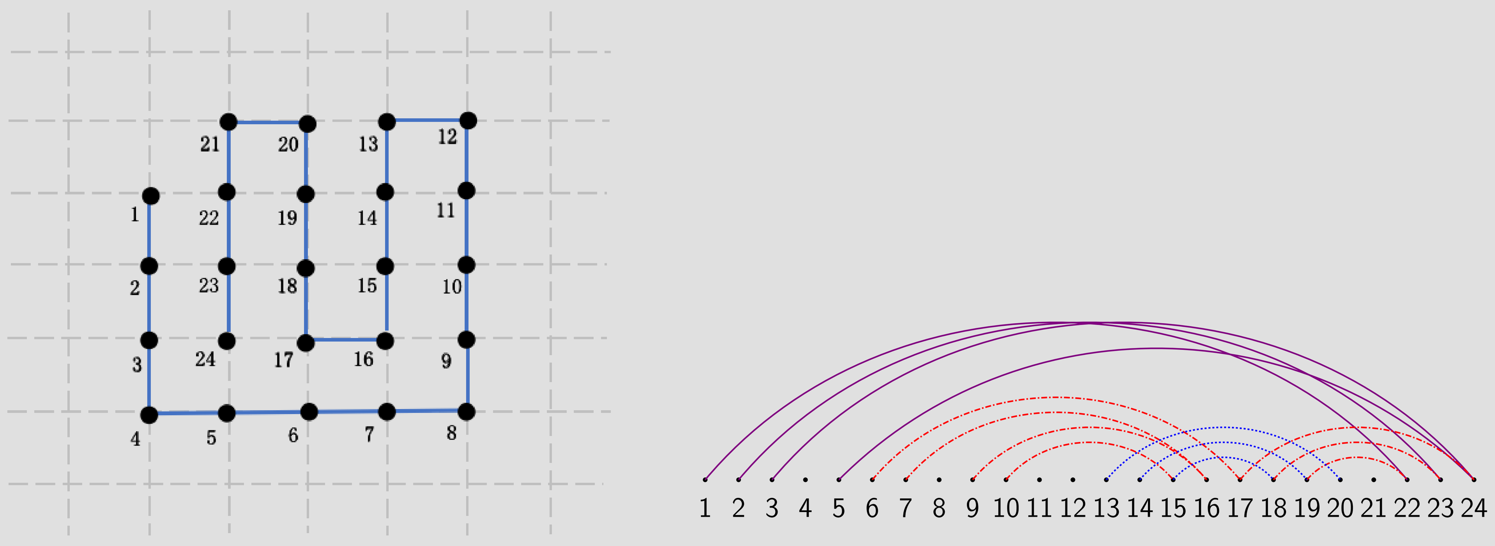


Figure: Protein fold in 2D square lattice (left) and its contact map (right).

- Goldman et al. (1999) showed that any protein contact map in the 2D square lattice can be decomposed into (at most) two stacks (noncrossing diagram) and one queue (nonnesting diagram).
- Istrail and Lam (2009) proposed the question in a review that "Are there generalizations of the Schmitt-Waterman counting formulas to stacks (in full generality) and to queues?"

General stacks and DLU-paths

General stacks:

- m-regular: arc length $\geq m$.
- d-contact: vertex degree $\leq d$.
- $S_{m,d}(x)$: generating function for m-regular d-contact stacks.

A DLU-path P of length dn:

- P is a Motzkin path from $(0,0)$ to $(dn,0)$ with steps in $\{D=(1,-1), L=(1,0), U=(1,1)\}$.
- It can be split into n segments $P=P_1P_2\cdots P_n$, such that each P_i is of length d and in the form of $D^{a_i}L^{b_i}U^{c_i}$ with $a_i+b_i+c_i=d$, $a_i,b_i,c_i\geq 0$ for $1\leq i\leq n$.
- m-regular: no subpath $UP'D$ with $P'\in\{L^0,\ldots,L^{(m-2)d}\}$.
- **Λ-avoiding (pattern avoidance)**: there is no subpath $UUP'DD$ with P' being empty or starting and ending at the same height $h\geq 2$ and never passing below the height h.
- $G_{m,d}^{\langle s,t\rangle}(x)$: generating function for partial DLU-paths from $(0,s)$ to (dn,t).

Bijection η: General stacks $\longleftrightarrow$ DLU-paths

- η: for each vertex i in a general stack S with $\mathrm{ld}(i)=a_i$, $\mathrm{rd}(i)=c_i$, it is mapped to a segment $P_i=D^{a_i}L^{d-a_i-c_i}U^{c_i}$ of the lattice path P.
- Inverse map: let a_i and c_i be the numbers of D and U steps in the segment P_i, which are used as the multiplicities of i in the multisets V_1 and V_2. From the two multisets $V_1=\{1^{a_1},2^{a_2},\ldots,n^{a_n}\}$ and $V_2=\{1^{c_1},2^{c_2},\ldots,n^{c_n}\}$, for the largest element $i\in V_2$ find $j\in V_1$ such that j is one of the smallest vertices in V_1 larger than i. Then add an arc (i,j) to S and delete a single i,j from V_2,V_1, respectively.

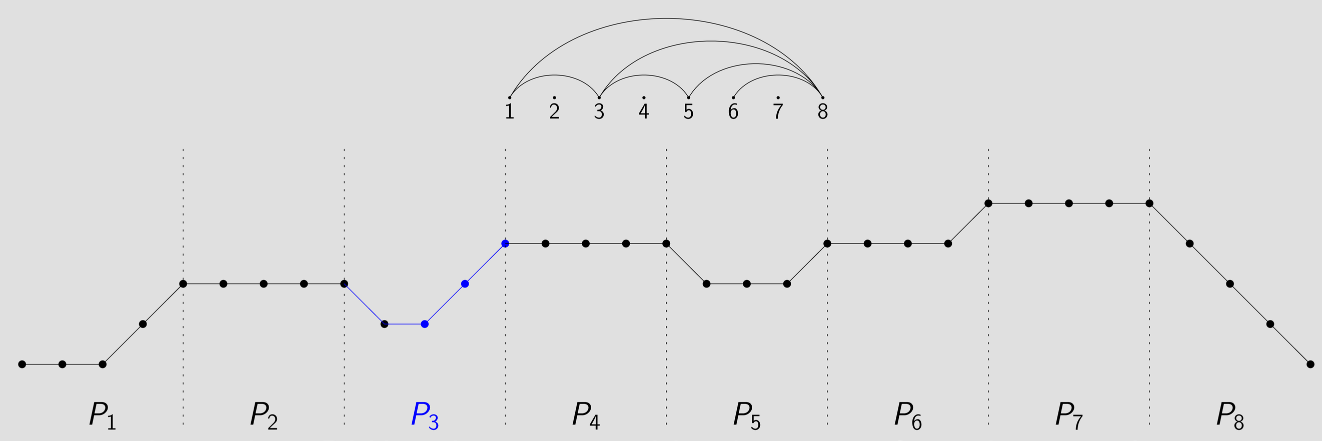


Figure: The 2-regular Λ-avoiding DLU-path constructed by the map η from the 2-regular 4-contact stack with $n=8$, $d=4$ and $V_1=\{3,5,8^4\}$, $V_2=\{1^2,3^2,5,6\}$.

Enumeration formulas for stacks

- $d=1, m=1$: $S_{1,1}(x)=\dfrac{1-x-\sqrt{1-2x-3x^2}}{2x^2}$ is well known as the generating function of Motzkin numbers.
- $d=1, m=2$: $S_{2,1}(x)=\dfrac{x^2-x+1-\sqrt{1+x(x^3-2x^2-x-2)}}{2x^2}$ is a classic result given by Waterman (1978) for the number of RNA secondary structures.
- $d=2$: the generating function of the **m-regular linear stacks** satisfies

$$\sum_{i=0}^{5} c_{m,2,i}(x)S_{m,2}^{i}(x)=0,$$

with explicit coefficients $c_{m,2,i}(x)$.

Unified system of equations for general stacks

The generating function $S_{m,d}(x)$ satisfies that $S_{m,d}(x^d)=G_{m,d}^{\langle 0,0\rangle}(x)$, and $G_{m,d}^{\langle 0,0\rangle}(x)$ can be determined by the following system of equations

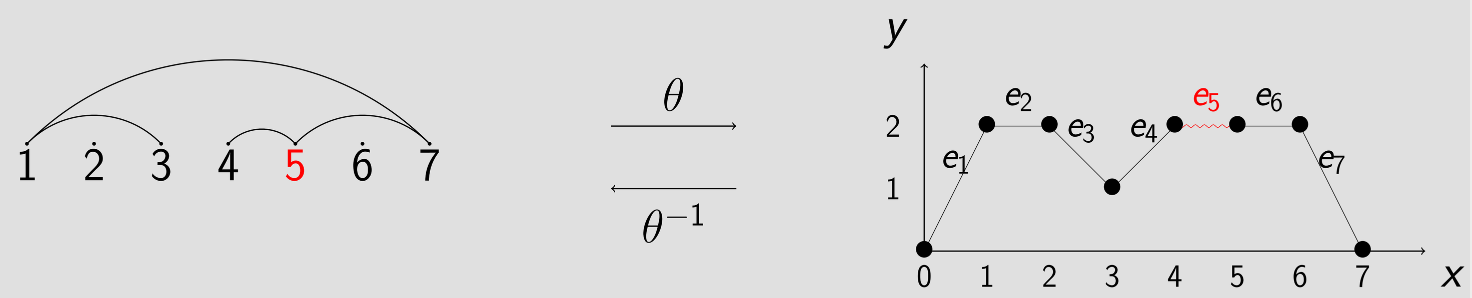
$$\begin{cases} G_{m,d}^{\langle 0,0\rangle}=1+x^d\displaystyle\sum_{i=0}^{d}G_{m,d}^{\langle i,0\rangle}, \\ G_{m,d}^{\langle s,t\rangle}=(-1)^t\delta_{s,t}C_{m,d}+\displaystyle\sum_{i=0}^{t-1}(-1)^{t-i+1}G_{m,d}^{\langle s-t+i,i\rangle} \\ \qquad +x^d\displaystyle\sum_{a=1}^{d}\left(\left((-1)^v\delta_{u,v}C_{m,d}+\sum_{j=0}^{v-1}(-1)^{v-j+1}G_{m,d}^{\langle u-v+j,j\rangle}\right)\cdot\sum_{c=0}^{d-a}G_{m,d}^{\langle c,t\rangle}\right), \\ \qquad\qquad 0\leq t\leq s\leq d-1,\ s\neq 0, \\ G_{m,d}^{\langle d,0\rangle}=x^d\displaystyle\sum_{a=1}^{d-1}\left(\left(\sum_{j=0}^{a-1}(-1)^{a-j+1}G_{m,d}^{\langle d-a+j,j\rangle}\right)\cdot\sum_{c=0}^{d-a}G_{m,d}^{\langle c,0\rangle}\right) \\ \qquad +x^d\left((-1)^dC_{m,d}+\displaystyle\sum_{j=0}^{d-1}(-1)^{d-j+1}G_{m,d}^{\langle j,j\rangle}\right)G_{m,d}^{\langle 0,0\rangle}, \end{cases}$$

where G, C are short for $G(x)$, $C(x)$, $u=\max\{s,a\}$, $v=\min\{s,a\}$, $\delta_{i,j}=1$ if $i=j$ or 0 if $i\neq j$, and $C_{m,d}(x)=\dfrac{x^{d(m-1)}-1}{x^d-1}$.

Generalized 2-Motzkin paths

For an m-regular linear stack $S_{m,h}$ with nesting number at most h, the Λ-avoiding m-regular generalized 2-Motzkin path P can be obtained by mapping each vertex in $S_{m,h}$ to a step in P following the correspondence rules given in the table:

vertex	$\mathrm{rd}(i)=1$ $\mathrm{ld}(i)=0$	$\mathrm{rd}(i)=2$ $\mathrm{ld}(i)=0$	$\mathrm{rd}(i)=0$ $\mathrm{ld}(i)=1$	$\mathrm{rd}(i)=0$ $\mathrm{ld}(i)=2$	$\mathrm{rd}(i)=0$ $\mathrm{ld}(i)=0$	$\mathrm{rd}(i)=1$ $\mathrm{ld}(i)=1$
step	$U^{[1]}=(1,1)$	$U^{[2]}=(1,2)$	$D^{[1]}=(1,-1)$	$D^{[2]}=(1,-2)$	$H=(1,0)$	$\widetilde{H}=\widetilde{(1,0)}$



From the above bijection and by using **the First Passage Decomposition method** of the lattice paths, we obtain a difference equation satisfied by the generating functions for the m-regular linear stacks with nesting number at most h.

Dyck paths with bounded height

Let $D_h(x,y)$ be the generating function of Dyck paths with p peaks and height $\leq h$.

- It satisfies the continued fraction:

$$D_h(x,y)=\frac{1}{1+x-xy-xD_{h-1}(x,y)},\quad h\geq 1,\quad D_0=1.$$

- When $h\geq 1$, $D_h(x,y)$ can be written as a rational function of the Chebyshev polynomials of the second kind

$$D_h(x,y)=\frac{1}{\sqrt{x}}\cdot\frac{U_{h-1}(t)-\sqrt{x}U_{h-2}(t)}{U_h(t)-\sqrt{x}U_{h-1}(t)}, \tag{2}$$

where

$$t=\frac{1+x-xy}{2\sqrt{x}}.$$

- It leads to a bijection between Dyck paths of height ≤ 2 with $p+1$ peaks and permutations avoiding patterns 123 and 132 with p descents.

Conclusions

- We established a unified combinatorial framework for counting general stacks with arbitrary vertex degree.
- The framework also extends to lattice paths with bounded heights, with applications to nesting numbers in biological structures.
- Extensions to more pattern avoiding lattice paths are possible.

References

[1] Guo, Q.H.; Jin, Y.L.; Sun, L.H.; Xu, S.N. "*Bijective enumeration of general stacks*". Adv. Appl. Math., 158, 102722, 2024.

[2] Li, C.Y.; Sun, L.H.: "*Enumeration of lattice paths with bounded height*". submitted, 2026.

[3] Cigler, J.; Krattenthaler, C.: "*Bounded Dyck paths, bounded alternating sequences, orthogonal polynomials, and reciprocity*". European J. Combin., 121, 103840, 2024.