JOURNAL OF COMPUTATIONAL BIOLOGY
Volume 15, Number 9, 2008

Pp. 1195–1208
DOI: 10.1089/cmb.2008.0051

Pseudoknot RNA Structures with Arc-Length ≥4

HILLARY S.W. HAN and CHRISTIAN M. REIDYS

ABSTRACT

In this article, we study k-noncrossing RNA structures with minimum arc-length 4 and at most $k-1$ mutually crossing bonds. Let $\mathsf{T}_k^{[4]}(n)$ denote the number of k-noncrossing RNA structures with arc-length ≥4 over n vertices. We (a) prove a functional equation for the generating function $\sum_{n\geq 0}\mathsf{T}_k^{[4]}(n)z^n$ and (b) derive for $4\leq k\leq 9$ the asymptotic formula $\mathsf{T}_k^{[4]}(n)\sim c_k\, n^{-((k-1)^2+(k-1)/2)}\,\gamma_k^{-n}$. Furthermore, we explicitly compute the exponential growth rates γ_k^{-1} and asymptotic formulas for $4\leq k\leq 9$.

Key words: exponential growth rate, generating function, k-noncrossing diagram, RNA pseudoknot structure, singularity analysis.

1. INTRODUCTION

THE RNA PSEUDOKNOT STRUCTURES (Mapping RNA, 2005; Westhof and Jaeger, 1992) are a reality. They occur in functional RNA (RNAseP [Loria and Pan, 1996]), ribosomal RNA (Konings and Gutell, 1995), and are conserved in the catalytic core of group I introns. Due to the crossings of arcs their theory differs considerably from RNA secondary structures. Pseudoknots are inherently noninductive and the standard dynamic programming folding paradigm employed for RNA secondary structures can only generate particular subclasses of pseudoknot structures (Rivas and Eddy, 1999). Recently the concept of k-noncrossing RNA structures has been introduced (Jin et al., 2008a). Here the idea is that the complexity of the structure is captured by an inherently "local" property: the maximal number of mutually crossing bonds. A structure is k-noncrossing, if there exists no k-set of mutually crossings arcs. The locality is in fact of central importance: point in case are RNA bisecondary structures introduced by Haslinger and Stadler (1999). These structures are constructed as superpositions of two RNA secondary structures and correspond to *planar* 3-noncrossing structures Jin et al. (2008a). The planarity property is clearly non-local and at present time the generating function for RNA bisecondary structures is not known.

A very intuitive approach to the k-noncrossing property of RNA molecules is their diagram representation (Haslinger and Stadler, 1999). It is obtained by drawing the nucleotide-labels $1,\ldots,n$ in increasing order in a horizontal line and drawing the arc-labels (i,j) in the upper half-plane, if and only if i and j are paired in the structure (Fig. 1). We call a diagram k-noncrossing, if it does not contain k mutually crossing arcs. The length of an arc (i,j) is given by $\lambda=j-i$ and a stack of length σ is a sequence of "parallel" arcs of the form $((i,j),(i+1,j-1),\ldots,(i+(\sigma-1),j-(\sigma-1)))$. A k-noncrossing RNA structure is a k-noncrossing diagram over $[n]$ having minimum arc-length $\lambda>1$. These structures have been studied

Center for Combinatorics, Nankai University, Tianjin, P.R. China.

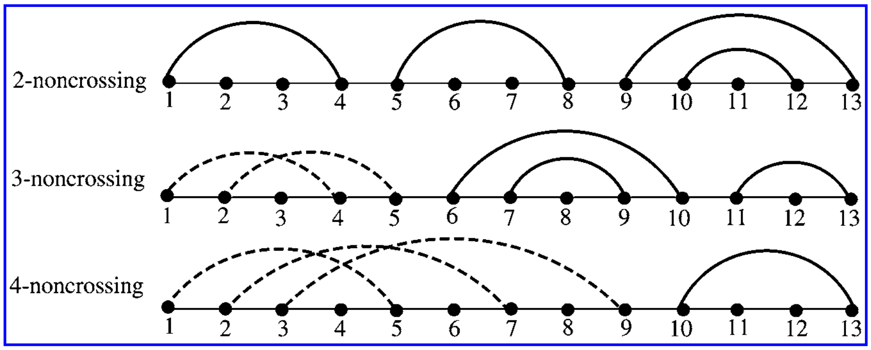

FIG. 1. k-Noncrossing structures: 2-, 3-, and 4-noncrossing structures (top to bottom). Maximal sets of mutually crossing arcs are dashed.

in Jin et al. (2008a) and Jin and Reidys (2008a) via a bijection into vacillating tableau in the context of tangled diagrams (Chen et al., 2008a). For the enumeration of structures with crossing arcs the tableaux-interpretation is non-optional. There is, to the best of our knowledge, no way to inductively construct k-noncrossing structures, despite the fact that they are D-finite.

For RNA secondary structures (2-noncrossing RNA structures), certain combinatorial restrictions, for instance minimum arc-length or stack-size are relatively straightforward to deal with. The combinatorics and prediction of RNA secondary structures has been pioneered by Waterman et al. in a series of excellent papers (Howell et al., 1980; Penner and Waterman, 1993; Waterman, 1978, 1979; Waterman and Schmitt, 1994). He proved for the number of RNA secondary structures of length n (arc-length ≥ 2), $\mathsf{T}_2^{[2]}(n)$, the fundamental recursion

$$\mathsf{T}_2^{[2]}(n) = \mathsf{T}_2^{[2]}(n-1) + \sum_{s=0}^{n-3} \mathsf{T}_2^{[2]}(n-2-s)\mathsf{T}_2^{[2]}(s), \tag{1.1}$$

where $\mathsf{T}_2^{[2]}(0) = \mathsf{T}_2^{[2]}(1) = \mathsf{T}_2^{[2]}(2) = 1$. Equation (1.1) is an immediate consequence considering secondary structures as peak-free Motzkin-paths, i.e., peak-free paths with *up*, *down*, and *horizontal* steps that stay in the upper halfplane, starting at the origin and end on the x-axis. The recursion is in particular the key for all asymptotic results since it immediately implies a functional equation for the corresponding generating function. This allows the application of Darboux-type theorems (Hofacker et al., 1998; Wong and Wyman, 1974). For the number of secondary structures with minimum arc-length λ, $\mathsf{T}_2^{[\lambda]}(n)$, it is straightforward to derive

$$\mathsf{T}_2^{[\lambda]}(n) = \mathsf{T}_2^{[\lambda]}(n-1) + \sum_{s=0}^{n-(\lambda+1)} \mathsf{T}_2^{[\lambda]}(n-2-s)\mathsf{T}_2^{[\lambda]}(s). \tag{1.2}$$

All asymptotic formulae for secondary structures are of the same type: a square root. In other words, the asymptotic behavior is determined by an algebraic branch singularity with the subexponential factor $n^{-\frac{3}{2}}$.

The situation changes for k-noncrossing RNA structures. A different approach has to be made, since given the lack of functional equations, Darboux-type theorems (Wong and Wyman, 1974) cannot be employed. The idea is to analyze the dominant singularities directly, using Hankel contours. Singularity analysis has been pioneered by Flajolet and Sedgewick (2008). Its basic idea is the construction of the "singular-analogue" of the Taylor-expansion. It can be shown that, under certain conditions, there exists an approximation, which is locally of the same order as the original function. The particular, local approximation allows then to derive the asymptotic form of the coefficients. In contrast to the subtraction of singularities-principle (Odlyzko, 1995), the only contributions to the contour integral come from segments close to the singularity. In our situation, all conditions for singularity analysis are satisfied, since the generating functions involved are D-finite (Stanley, 1980; Zeilberger, 1990) and D-finite functions have an analytic continuation into any simply-connected domain containing zero. Our approach also works for tangled diagrams (Chen et al., 2008b), which represent the combinatorial framework for RNA tertiary interactions. Our analysis confirms that the particular singularity-type of the generating function of k-noncrossing RNA structures depends solely on the crossing number (Jin and Reidys, 2008a, 2008b). While the location of the singularity shifts as a function of the arc-length, all subexponential factors remain the same. Furthermore, an interesting feature is the appearance of logarithms for $k \equiv 1 \bmod 2$ in the singular expansion.

Due to biophysical constraints a minimum arc-length of four can be assumed for minimum free energy RNA structures. The key objective of this article is to derive and analyze the generating function for k-noncrossing RNA structures with minimum arc-length 4 (Table 1). Based on our results, the next step is to compute the subset of canonical structures, i.e. the subset of structures with arc-length ≥ 4, having no isolated arcs. While it is straightforward to obtain Equation (1.2) from Equation (1.1) considerable complication arises, when considering k-noncrossing structures with arc-length > 3. To understand why, one observes that the number of ways to place 3-arcs satisfies a new type of recursion; see Equation (3.6). As a result and in contrast to k-noncrossing structures with minimum arc-length $\lambda \leq 3$ the generating function $\sum_{n\geq 0} \mathsf{T}_k^{[4]}(n)\, z^n$ turns out to be a sum of two power series (Theorem 2). The exponential growth rate can easily be computed via the formula given in Theorem 3 (Table 2 and Fig. 3).

The article is organized as follows: in Section 2, we provide the background on the methods used in this article. In Section 3, we prove a functional equation relating RNA structures to k-noncrossing matchings. We then study the singularity of the generating function and obtain the asymptotic formula in Section 4. Finally, in Section 5, we detail some key ideas instrumental for the proof of Theorem 2.

2. PRELIMINARIES

In this section, we provide some background on the generating functions of k-noncrossing matchings (Chen et al., 2007; Jin et al., 2008b) and k-noncrossing RNA structures (Jin et al., 2008a; Jin and Reidys, 2008a). We denote the numbers of k-noncrossing matchings of length $2n$ and RNA structures with arc-length $\geq\lambda$ of length n by $f_k(2n)$ and $\mathsf{T}_k^{[\lambda]}(n)$, respectively. The former corresponds to k-noncrossing diagrams without isolated points and the latter to k-noncrossing diagrams with arc-length $\geq\lambda$. Furthermore, let $\mathsf{T}_k^{[\lambda]}(n,\ell)$ denote the number of k-noncrossing RNA structures with arc-length $\geq\lambda$ having exactly ℓ isolated points and $\mathsf{M}_k(n)$ denotes the number of partial matchings, or equivalently the number of k-noncrossing diagrams over $[n]$ (i.e., with isolated points and minimum arc-length 1). Pfringsheim's Theorem (Titchmarsh, 1939) guarantees the existence of a positive real, dominant singularity of $\sum_{n\geq 0} \mathsf{M}_k(n)\, z^n$ which we denote by μ_k. In order to get some intuition about various types of diagrams involved, see Figure 2.

2.1. *k-Noncrossing matchings*

Our main objective is to discuss some basic properties of $f_k(2n)$ and to give an asymptotic formula. Let us recall that a power series $u(x)$ is called D-finite over the function field $K(x)$ if $\dim\langle u, u', \ldots\rangle_{K(x)} < \infty$ (Stanley, 1980). The generating function of k-noncrossing matchings satisfies the following identity due to Grabiner and Magyar (1993)

$$\sum_{n\geq 0} f_k(2n)\cdot \frac{z^{2n}}{(2n)!} = \det[I_{i-j}(2z) - I_{i+j}(2z)]|_{i,j=1}^{k-1}, \tag{2.1}$$

where

$$I_r(2z) = \sum_{j\geq 0} \frac{z^{2j+r}}{j!(r+j)!} \tag{2.2}$$

denotes the hyperbolic Bessel function of the first kind of order r. Equation (2.1) allows us to conclude that

$$F_k(z) = \sum_{n\geq 0} f_k(2n) z^{2n} \tag{2.3}$$

TABLE 1. THE FIRST 15 NUMBERS OF 4-NONCROSSING RNA STRUCTURES WITH ARC-LENGTH ≥4

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*
$\mathsf{T}_4^{[4]}(n)$	1	1	1	1	2	5	15	51	179	647	2397	9,081	35,181	139,307	563,218
$\mathsf{T}_5^{[4]}(n)$	1	1	1	1	2	5	15	52	188	703	2704	10,684	43,376	180,971	775,016

TABLE 2. EXPONENTIAL GROWTH RATES AND ASYMPTOTIC FORMULAS FOR k-NONCROSSING RNA STRUCTURES WITH MINIMUM ARC-LENGTH ≥4

k	*4*	*5*	*6*	*7*	*8*
γ_k^{-1}	6.52900	8.64830	10.71759	12.76349	14.79631
$\mathsf{T}_k^{[4]}(n)$	$c_4 n^{-\frac{21}{2}}(\gamma_4^{-1})^n$	$c_5 n^{-18}(\gamma_5^{-1})^n$	$c_6 n^{-\frac{55}{2}}(\gamma_6^{-1})^n$	$c_7 n^{-39}(\gamma_7^{-1})^n$	$c_8 n^{-\frac{105}{2}}(\gamma_8^{-1})^n$

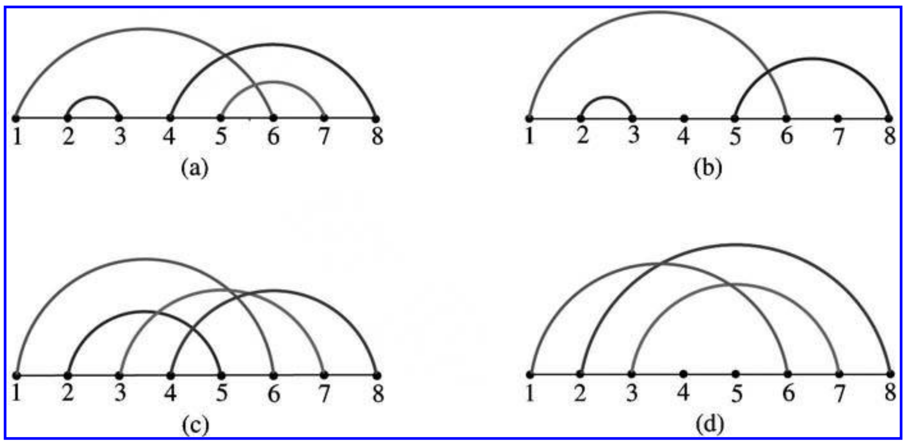

FIG. 2. Basic diagram types. **(a)** 3-noncrossing matching (no isolated points). **(b)** 3-noncrossing partial matching (isolated points 4 and 7). **(c)** 4-noncrossing RNA structure with arc-length ≥3. **(d)** 3-noncrossing RNA structure with arc-length ≥4.

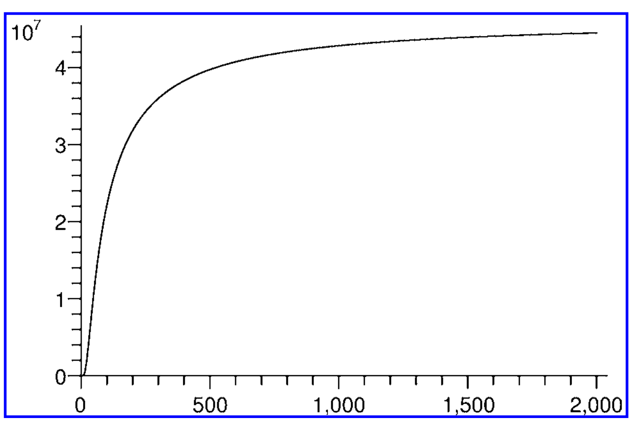

FIG. 3. The ratio $r(n) = \mathsf{T}_4^{[4]}(n)/(n^{-21/2}\gamma_4^{-n})$ as a function of n. The curve shows that the asymptotic approximation is valid as $r(n) \sim c_4 \approx 4.4509 \times 10^7$.

is D-finite. Indeed, the hyperbolic Bessel function (Grabiner and Magyar, 1993) itself is D-finite and D-finite functions form an algebra closed under taking Hadamard products (Stanley, 1980). Therefore, D-finiteness of $F_k(z)$ follows from Equation (2.1). However, beyond the cases $k = 2$ and $k = 3$, Equation (2.1) does not give directly explicit formulas for $f_k(2n)$ or $F_k(z)$. For small k-values, asymptotic formulas can be obtained using the approximation of the Bessel function

$$I_m(z) = \frac{e^z}{\sqrt{2\pi z}}\left(\sum_{h=0}^{H-1}\frac{(-1)^h}{h!8^h}\prod_{t=1}^{h}(4m^2-(2t-1)^2)z^{-h} + O(|z|^{-H})\right) \tag{2.4}$$

which holds for $-\frac{\pi}{2} < \arg(z) < \frac{\pi}{2}$ (Abramowitz and Stegun, 1964). For arbitrary k, systematic analysis of the determinant $\det[I_{i-j}(2x) - I_{i+j}(2x)]|_{i,j=1}^{k-1}$ by Jin et al. (2008b) shows for arbitrary k

$$f_k(2n) \sim c_k\, n^{-((k-1)^2+(k-1)/2)}\,(2(k-1))^{2n}, \qquad c_k > 0. \tag{2.5}$$

In the following, we shall denote the dominant singularities of $F_k(z)$ by $\rho_k = \frac{1}{2(k-1)}$ and $-\rho_k$, respectively.

2.2. *k-Noncrossing RNA structures*

k-Noncrossing RNA structures are k-noncrossing diagrams satisfying specific arc-length conditions. The latter induce asymmetries (for instance, 1-arcs are not preserved) which prohibit enumeration using Gessel and Zeilberger's (1992) reflection-principle directly (the reflection principle implies Equation (2.1)). For any $k \geq 2$ the numbers of k-noncrossing RNA structures with minimum arc-length ≥ 2 are given by Jin et al. (2008a)

$$\mathsf{T}_k^{[2]}(n) = \sum_{b=0}^{\lfloor n/2\rfloor}(-1)^b\binom{n-b}{b}\mathsf{M}_k(n-2b) \tag{2.6}$$

and we have (Jin and Reidys, 2008a)

$$\mathsf{T}_k^{[2]}(n) \sim c_k^{[2]}\, n^{-((k-1)^2+(k-1)/2)}\,(\gamma_k^{[2]})^{-n}, \qquad c_k^{[2]} > 0, \tag{2.7}$$

where $\gamma_k^{[2]}$ is the unique solution of minimal modulus of $\frac{z}{z^2-z+1} = \rho_k = \frac{1}{2(k-1)}$. For k-noncrossing RNA structures with arc-length ≥ 3, we have according to Jin et al. (2008a)

$$\forall\, k > 2; \qquad \mathsf{T}_k^{[3]}(n) = \sum_{b\leq\lfloor\frac{n}{2}\rfloor}(-1)^b\lambda(n,b)\,\mathsf{M}_k(n-2b), \tag{2.8}$$

where $\lambda(n,b)$ denotes the number of ways selecting b arcs of length ≤ 2 over n vertices. The nonexplicit terms $\lambda(n,b)$ vanish in the functional equation $\sum_{n\geq 0}\mathsf{T}_k^{[3]}(n)\,z^n$ which equals the following:

$$\frac{1}{1-z+z^2+z^3-z^4}\sum_{n\geq 0}f_k(2n)\left(\frac{z-z^3}{1-z+z^2+z^3-z^4}\right)^{2n}. \tag{2.9}$$

Singularity analysis based on Equation (2.9) eventually allows to derive the asymptotic formula

$$\mathsf{T}_k^{[3]}(n) \sim c_k^{[3]}\, n^{-((k-1)^2+(k-1)/2)}\,(\gamma_k^{[3]})^{-n}, \qquad c_k^{[3]} > 0, \tag{2.10}$$

where $\gamma_k^{[3]}$ denotes the unique, minimal positive real solution of $\frac{z-z^3}{1-z+z^2+z^3-z^4} = \rho_k$.

2.3. Singularity analysis

Pfringsheim's Theorem (Titchmarsh, 1939) guarantees that each power series with positive coefficients has a positive real dominant singularity. This singularity plays a key role for the asymptotics of the coefficients. In the proof of Theorem 3, it will be important to deduce relations between the coefficients from functional equations of generating functions. The class of theorems that deal with such deductions are called transfer-theorems (Flajolet and Sedgewick, 2008). One key ingredient in this framework is a specific domain in which the functions in question are analytic, which is "slightly" bigger than their respective radius of convergence. It is tailored for extracting the coefficients via Cauchy's integral formula. Details on the method can be found in Flajolet and Sedgewick (2008) and Stanley (1980). In case of D-finite functions, we have analytic continuation in any simply-connected domain containing zero (Wasow, 1987) and all prerequisits of singularity analysis are met. We use the notation

$$\{f(z) = O(g(z)) \text{ as } z \to \rho\} \quad \Longleftrightarrow \quad \left\{\frac{f(z)}{g(z)} \text{ is bounded as } z \to \rho\right\} \tag{2.11}$$

The key result used in Theorem 3 is as follows:

Theorem 1 (Flajolet and Sedgewick, 2008). *Let $f(z), g(z)$ be D-finite functions with unique dominant singularity ρ and suppose*

$$f(z) = O(g(z)) \quad as \ z \to \rho. \tag{2.12}$$

Then we have

$$[z^n]f(z) = C\left(1 - O\left(\frac{1}{n}\right)\right)[z^n]g(z) \tag{2.13}$$

where C is a constant and $[z^n]h(z)$ denotes the nth coefficient of the power series $h(z)$ at $z = 0$.

3. THE GENERATING FUNCTION

In this section, we compute the generating function of $\mathsf{T}_k^{[4]}(n)$, the number of k-noncrossing RNA structures with arc-length ≥ 4. Our first result is a technical lemma which is instrumental in the proof of Theorem 2 below. The proof of the lemma given below is new and uses integral representations (Egorychev, 1984) instead of dealing with the combinatorial coefficients directly.

Lemma 1. *Let z be an indeterminate over $\mathbb{C}$. Then we have the identity of power series*

$$\forall\ |z| < \mu_k; \qquad \sum_{n\geq 0} \mathsf{M}_k(n)\, z^n = \left(\frac{1}{1-z}\right)\sum_{n\geq 0} f_k(2n)\left(\frac{z}{1-z}\right)^{2n}. \tag{3.1}$$

Proof. Expressing the combinatorial terms by contour integrals (Egorychev, 1984), we obtain

$$\binom{n}{2m} = \frac{1}{2\pi i}\oint_{|u|=\alpha}(1+u)^n u^{-2m-1}du \qquad f_k(2m) = \frac{1}{2\pi i}\oint_{|v|=\beta}F_k(v)v^{-2m-1}dv \tag{3.2}$$

where α, β are arbitrary small positive numbers. We derive

$$\begin{aligned}\mathsf{M}_k(n) &= \frac{1}{(2\pi i)^2}\sum_m \oint_{|u|=\alpha,|v|=\beta}(1+u)^n u^{-2m-1}F_k(v)v^{-2m-1}dudv\\ &= \frac{1}{(2\pi i)^2}\oint_{|u|=\alpha,|v|=\beta}(1+u)^n\frac{uv}{(uv)^2-1}F_k(v)dudv\end{aligned}$$

and furthermore

$$\mathsf{M}_k(n) = \frac{1}{(2\pi i)^2} \oint_{|v|=\beta} F_k(v) v^{-1} \left[\oint_{|u|=\alpha} \frac{(1+u)^n u}{(u+\frac{1}{v})(u-\frac{1}{v})} du \right] dv.$$

Since $u = \frac{1}{v}$ and $u = -\frac{1}{v}$ are the only singularities (poles) enclosed by the particular contour, Equation (3.1) implies

$$\oint_{|u|=\alpha} \frac{(1+u)^n u}{(u+\frac{1}{v})(u-\frac{1}{v})} du = 2\pi i \left[\frac{(1+u)^n u}{u-\frac{1}{v}} |_{u=-\frac{1}{v}} + \frac{(1+u)^n u}{u+\frac{1}{v}} |_{u=\frac{1}{v}} \right]$$

$$= \pi i \left(\left[1 - \frac{1}{v}\right]^n + \left[1 + \frac{1}{v}\right]^n \right).$$

Therefore, for $|z| < \mu_k$

$$\sum_{n \geq 0} \mathsf{M}_k(n) z^n = \frac{1}{4\pi i} \sum_{n \geq 0} \oint_{|v|=\beta} F_k(v) v^{-1} \left(\left[1 - \frac{1}{v}\right]^n + \left[1 + \frac{1}{v}\right]^n \right) z^n dv$$

$$= \frac{1}{4\pi i} \oint_{|v|=\beta} F_k(v) \frac{1}{v-(v-1)z} dv + \frac{1}{4\pi i} \oint_{|v|=\beta} F_k(v) \frac{1}{v-(v+1)z} dv.$$

The first integrand has its unique pole at $v = -\frac{z}{1-z}$ and the second at $v = \frac{z}{1-z}$, respectively:

$$\frac{1}{v-(v-1)z} = \frac{1}{v+\frac{z}{1-z}} \frac{1}{1-z} \quad \text{and} \quad \frac{1}{v-(v+1)z} = \frac{1}{v-\frac{z}{1-z}} \frac{1}{1-z}.$$

In view of $F_k(z) = F_k(-z)$ we derive

$$\sum_{n \geq 0} \mathsf{M}_k(n) z^n = \frac{1}{1-z} \left[\frac{1}{2} F_k\left(-\frac{z}{1-z}\right) + \frac{1}{2} F_k\left(\frac{z}{1-z}\right) \right] = \frac{1}{1-z} F_k\left(\frac{z}{1-z}\right),$$

whence the lemma. ■

Before we state the main result of this section, let us introduce some notation. We set

$$u(z) = \sqrt{1 + 4z - 4z^2 - 6z^3 + 4z^4 + z^6} \tag{3.3}$$

$$f_j(z) = -\frac{-2z^2 + z^3 - 1 + (-1)^j \, u(z)}{2(1 - 2z - z^2 + z^4)}. \tag{3.4}$$

Note that $f_j(z)$ is an algebraic function over the function field $K(z)$, i.e., there exists a polynomial with coefficients being polynomials in z for which $f_j(z)$ is a root. This fact will be important when computing the subexponential factors of the asymptotic formula for $\mathsf{T}_k^{[4]}(n)$. We can now compute the generating function $\sum_{n \geq 0} \mathsf{T}_k^{[4]}(n) \, z^n$.

Theorem 2. *Let k be a positive integer, $k > 3$ and $f_1(z)$ and $f_2(z)$ be given by Equation (3.4). Then we have*

$$\mathsf{T}_k^{[4]}(n) = \sum_{b \leq \lfloor \frac{n}{2} \rfloor} (-1)^b \, \lambda(n, b) \, \mathsf{M}_k(n - 2b) \tag{3.5}$$

and $\lambda(n,b)$ *satisfies the recurrence*

$$\begin{aligned}\lambda(n,b) &= \lambda(n-1,b)+\lambda(n-4,b-2)+\lambda(n-5,b-2)+\lambda(n-6,b-3)\\ &\quad+\sum_{i=1}^{b}[\lambda(n-2i,b-i)+2\lambda(n-2i-1,b-i)+\lambda(n-2i-2,b-i)]\\ &\quad-\lambda(n-3,b-1),\end{aligned} \tag{3.6}$$

where $\lambda(n,0)=1$*,* $\lambda(n,1)=3n-6$ *and* $n\geq 2b$*. Furthermore*

$$\begin{aligned}\sum_{n\geq 0}\mathsf{T}_k^{[4]}(n)\,z^n &= \frac{F_1(-z^2)}{1-zf_1(-z^2)}\sum_{n\geq 0}f_k(2n)\left(\frac{z\,f_1(-z^2)}{1-zf_1(-z^2)}\right)^{2n}\\ &\quad+\frac{F_2(-z^2)}{1-zf_2(-z^2)}\sum_{n\geq 0}f_k(2n)\left(\frac{z\,f_2(-z^2)}{1-zf_2(-z^2)}\right)^{2n}.\end{aligned} \tag{3.7}$$

Proof. *Claim* 1. Let $\lambda(n,b)$ denote the number of ways to place b arcs of length <4 over $[n]$. Then we have

$$\mathsf{T}_k^{[4]}(n) = \sum_{b\leq\lfloor\frac{n}{2}\rfloor}(-1)^b\,\lambda(n,b)\,\mathsf{M}_k(n-2b) \tag{3.8}$$

and $\lambda(n,b)$ satisfies the recursion given via Equation (3.6). The proof of Claim 1 is analogous to the proof of Theorem 5 in Jin et al. (2008a). In order to keep the article self-contained, we present it in Section 5.

The idea is to use Equation (3.8) in order to relate $\sum_{n\geq 0}\mathsf{T}_k^{[4]}(n)\,z^n$ to the power series $\sum_{n\geq 0}\mathsf{M}_k(n)\,z^n$. For this purpose, we compute

$$\begin{aligned}\sum_{n\geq 0}\mathsf{T}_k^{[4]}(n)z^n &= \sum_{n\geq 0}\sum_{2b\leq n}(-1)^b\lambda(n,b)\sum_{m=2b}^{n}\binom{n-2b}{m-2b}f_k(m-2b,0)\,z^n\\ &= \sum_{b\geq 0}(-1)^b z^{2b}\sum_{n\geq 2b}\lambda(n,b)\,\mathsf{M}_k(n-2b)z^{n-2b}\\ &= \sum_{b\geq 0}(-1)^b z^{2b}\sum_{n\geq 0}\lambda(n+2b,b)\,\mathsf{M}_k(n)\,z^n.\end{aligned}$$

Interchanging the summations w.r.t. b and n, we arrive at

$$\sum_{n\geq 0}\mathsf{T}_k^{[4]}(n)z^n = \sum_{n\geq 0}\left[\sum_{b\geq 0}(-1)^b z^{2b}\lambda(n+2b,b)\right]\mathsf{M}_k(n)\,z^n. \tag{3.9}$$

Now we employ the recursion formula for $\lambda(n,b)$, Equation (3.6) in order to derive a functional equation for the generating function

$$\varphi_n(z) = \sum_{b\geq 0}\lambda(n+2b,b)z^b. \tag{3.10}$$

Multiplying in Equation (3.6) with z^b and taking the summation over all b ranging from 0 to $\lfloor n/2\rfloor$ implies for $\varphi_n(z)$, $n=1,2\ldots$, we obtain

$$\left(1-z^2-z^3-\frac{z}{1-z}\right)\varphi_n(z) = \left(z^2+\frac{z^2+1}{1-z}\right)\varphi_{n-1}(z)+\left(\frac{z}{1-z}\right)\varphi_{n-2}(z). \tag{3.11}$$

We make the Ansatz

$$f(x, y) = \sum_{n \geq 0} \sum_{j \leq \frac{n}{2}} \lambda(n, j) x^j \, \frac{y^n}{n!} = \sum_{n \geq 0} \varphi_n(x) \, \frac{y^n}{n!}. \tag{3.12}$$

Multiplying in Equation (3.11) with $\frac{y^n}{n!}$ and taking the summation over all $n \geq 0$ leads to the partial differential equation

$$\left(1 - x^2 - x^3 - \frac{x}{1-x}\right) \frac{\partial^2 f(x, y)}{\partial y^2} = \left(x^2 + \frac{x^2+1}{1-x}\right) \frac{\partial f(x, y)}{\partial y} + \left(\frac{x}{1-x}\right) f(x, y). \tag{3.13}$$

The general solution of Equation (3.13) can be computed by MAPLE and is given by

$$\begin{aligned} f(x, y) &= F_1(x) \exp(f_1(x) \cdot y) + F_2(x) \exp(f_2(x) \cdot y) \\ &= \sum_{n \geq 0} [F_1(x)\, f_1(x)^n + F_2(x)\, f_2(x)^n] \frac{y^n}{n!}, \end{aligned}$$

where $F_1(x)$, $F_2(x)$ are arbitrary functions and

$$f_1(x) = \frac{2x^2 - x^3 + 1 + u(x)}{2(1 - 2x - x^2 + x^4)}, \qquad f_2(x) = \frac{2x^2 - x^3 + 1 - u(x)}{2(1 - 2x - x^2 + x^4)}. \tag{3.14}$$

By definition, we have $f(x, y) = \sum_{n \geq 0} \varphi_n(x) \cdot \frac{y^n}{n!}$ and

$$\varphi_n(x) = F_1(x)(f_1(x))^n + F_2(x)(f_2(x))^n. \tag{3.15}$$

In order to solve Equation (3.15), it remains to compute $F_1(x)$ and $F_2(x)$. The key information lies in the initial conditions for $f(x, y)$ and $\varphi_n(x)$. Explicitly, we have $f(x, 0) = 1$ and $\varphi_1(x) = \lambda(1, 0)\, x^0 = 1$, which implies

$$\begin{aligned} F_1(x) + F_2(x) &= 1 \\ F_1(x) f_1(x) + F_2(x) f_2(x) &= 1. \end{aligned}$$

Accordingly, we obtain

$$F_1(x) = \frac{f_2(x) - 1}{f_2(x) - f_1(x)} \quad \text{and} \quad F_2(x) = \frac{f_1(x) - 1}{f_1(x) - f_2(x)}. \tag{3.16}$$

In view of $\varphi_n(-z^2) = \sum_{b \geq 0} \lambda(n + 2b, b)(-1)^b z^{2b}$, we can express $\sum_{n \geq 0} \mathsf{T}_k^{[4]}(n) z^n$ as follows:

$$\begin{aligned} \sum_{n \geq 0} \mathsf{T}_k^{[4]}(n) z^n &= \sum_{n \geq 0} \varphi_n(-z^2)\, \mathsf{M}_k(n)\, z^n \\ &= F_1(-z^2) \sum_{n \geq 0} \mathsf{M}_k(n)\, (f_1(-z^2) z)^n + F_2(-z^2) \sum_{n \geq 0} \mathsf{M}_k(n)\, (f_2(-z^2) z)^n. \end{aligned}$$

Now we use Lemma 1:

$$\sum_{n \geq 0} \mathsf{M}_k(n)\, z^n = \left(\frac{1}{1-z}\right) \sum_{n \geq 0} f_k(2n) \left(\frac{z}{1-z}\right)^{2n},$$

which allows us to express $\sum_{n\geq 0} \mathsf{T}_k^{[4]}(n) z^n$ via $\sum_{n\geq 0} f_k(2n)\, z^{2n}$

$$\sum_{n\geq 0} \mathsf{T}_k^{[4]}(n)\, z^n = \frac{F_1(-z^2)}{1-zf_1(-z^2)} \sum_{n\geq 0} f_k(2n) \left(\frac{zf_1(-z^2)}{1-zf_1(-z^2)}\right)^{2n} + \frac{F_2(-z^2)}{1-zf_2(-z^2)} \sum_{n\geq 0} f_k(2n) \left(\frac{zf_2(-z^2)}{1-zf_2(-z^2)}\right)^{2n}. \qquad \blacksquare$$

4. ASYMPTOTICS OF RNA PSEUDOKNOT STRUCTURES WITH ARC-LENGTH ≥4

We set

$$\vartheta_1(z) = \frac{z\, f_1(-z^2)}{1-zf_1(-z^2)} \tag{4.1}$$

$$\vartheta_2(z) = \frac{z\, f_2(-z^2)}{1-zf_2(-z^2)}. \tag{4.2}$$

Note that $\vartheta_1(z)$ and $\vartheta_2(z)$ are algebraic functions over the function field $K(z)$.

Theorem 3. *Let $k > 3$ be a positive integer and ρ_k, γ_k denote the positive real singularities of $F_k(z) = \sum_{n\geq 0} f_k(2n)z^{2n}$ and $\sum_{n\geq 0} \mathsf{T}_k^{[4]}(n)\, z^n$, respectively. Then the number of k-noncrossing RNA structures with arc-length ≥ 4 is for $k \leq 9$ asymptotically given by*

$$\mathsf{T}_k^{[4]}(n) \sim c_k\, n^{-((k-1)^2+(k-1)/2)}\, (\gamma_k^{-1})^n, \tag{4.3}$$

where γ_k is the unique positive, real solution of the equation $\vartheta_1(z) = \rho_k$.

Proof. According to Theorem 2 we have the functional equation

$$\sum_{n\geq 0} \mathsf{T}_k^{[4]}(n)\, z^n = \frac{F_1(-z^2)}{1-zf_1(-z^2)} \underbrace{\sum_{n\geq 0} f_k(2n) \left(\frac{z\, f_1(-z^2)}{1-zf_1(-z^2)}\right)^{2n}}_{F_k(\vartheta_1(z))} + \frac{F_2(-z^2)}{1-zf_2(-z^2)} \underbrace{\sum_{n\geq 0} f_k(2n) \left(\frac{z\, f_2(-z^2)}{1-zf_2(-z^2)}\right)^{2n}}_{F_k(\vartheta_2(z))}.$$

We consider the functions $\vartheta_1(z)$, $\vartheta_2(z)$ given by Equation (4.1) and Equation (4.2). The mappings $x \mapsto \vartheta_1(x)$ and $x \mapsto \vartheta_2(x)$ are strictly monotone and $\vartheta_1(x) > \vartheta_2(x)$ for $\vartheta_1(x) \in]0, \frac{1}{5}]$. Furthermore, we have $\rho_k < \rho_4 = \frac{1}{6}$, for $k > 4$. We can conclude from this that the real, positive dominant singularity, γ_k, of $\sum_{n\geq 0} \mathsf{T}_k^{[4]}(n)\, z^n$, whose existence is guaranteed by Pfringsheim's Theorem (Titchmarsh, 1939), satisfies

$$\vartheta_1(\gamma_k) = \rho_k. \tag{4.4}$$

Being a determinant of Bessel functions (Grabiner and Magyar, 1993), $F_k(z)$ is D-finite. Moreover $\vartheta_1(z)$ and $\vartheta_2(z)$ are algebraic over $K(z)$, analytic for $|z| < \delta$, where $\gamma_k < \delta$ and satisfy $\vartheta_1(0) = \vartheta_2(0) = 0$. Therefore, the composition $F_k(\vartheta_i(z))$, $i = 1, 2$, is D-finite (Stanley, 1980) and $F_k(\vartheta_1(z))$ and $F_k(\vartheta_2(z))$ have singular expansions, respectively. We further observe that neither $\frac{F_1(-z^2)}{1-zf_1(-z^2)}$ nor $\frac{F_2(-z^2)}{1-zf_2(-z^2)}$ have a singularity ζ with $|\zeta| \leq \gamma_k$. Hence, if ζ is a dominant singularity of $\sum_n \mathsf{T}_k^{[4]}(n)\, z^n$, then it is necessarily a singularity of $F_k(\vartheta_1(z))$ or $F_k(\vartheta_2(z))$. As for singularities of $F_k(\vartheta_1(z))$ and $F_k(\vartheta_2(z))$, we consider for $k \leq 9$ the ODE satisfied by $F_k(z)$:

$$q_{0,k}(z)\frac{d^e}{dz^e}F_k(z) + q_{1,k}(z)\frac{d^{e-1}}{dz^{e-1}}F_k(z) + \cdots + q_{e,k}(z)F_k(z) = 0, \tag{4.5}$$

where $q_{j,k}(z)$ are polynomials. The key point is now that any dominant singularity of $F_k(z)$ is contained in the set of roots of $q_{0,k}(z)$ (Stanley, 1980). Computing the ODEs for $4 \leq k \leq 9$ we can therefore conclude that $F_k(z)$ has only the two dominant singularities ρ_k and $-\rho_k$. Let $S = \{\zeta \mid \vartheta_1(\zeta) = \rho_k \text{ or } \vartheta_2(\zeta) = -\rho_k\}$. Then γ_k is the unique S-element of minimal modulus. We can draw two conclusions: first,

$$[z^n]\,\mathsf{T}_k^{[4]}(z) \sim c_k\,[z^n]\,F_k(\vartheta_1(z)) \quad \text{for some } c_k > 0 \tag{4.6}$$

and second, γ_k is the unique dominant singularity of $\sum_n \mathsf{T}_k^{[4]}(n)\,z^n$. In view of Equation (4.6), it thus remains to analyze the subexponential factors of the singular expansion of $F_k(\vartheta_1(z))$ at $z = \gamma_k$. Since $\vartheta_1(z)$ is regular at γ_k, we are given the supercritical case of singularity analysis (Flajolet and Sedgewick, 2008). In the supercritical case, the subexponential factors of the compositum, $F_k(\vartheta_1(z))$ coincide with those of the outer function, $F_k(z)$. According to Jin et al. (2008b), we have for arbitrary k

$$f_k(2n) \sim n^{-((k-1)^2+\frac{k-1}{2})}\,(\rho_k{}^{-1})^{2n} \tag{4.7}$$

and therefore the subexponential factors of $F_k(z) = \sum_{n\geq 0} f_k(2n)z^{2n}$ coincide with those of $F_k(\vartheta_1(z))$, i.e., we have

$$\mathsf{T}_k^{[4]}(n) \sim c_k\,n^{-((k-1)^2+\frac{k-1}{2})}\,(\gamma_k{}^{-1})^n \tag{4.8}$$

proving the theorem. ■

5. PROOF OF CLAIM 1

We recall that the numbers of k-noncrossing matchings and RNA structures with arc-length $\geq\lambda$ are denoted by $f_k(2n)$ and $\mathsf{T}_k^{[\lambda]}(n)$, respectively. Furthermore, $\mathsf{T}_k^{[\lambda]}(n,\ell)$ denotes the number of k-noncrossing RNA structures with arc-length $\geq\lambda$ having exactly ℓ isolated points, and let $f_k(m,\ell)$ denote the number of k-noncrossing diagrams with ℓ isolated points over m vertices. Let $\mathcal{G}_{n,k}(\ell, j_1, j_2, j_3)$ be the set of all k-noncrossing diagrams having exactly ℓ isolated points and exactly j_1 1-arcs, j_2 2-arcs, and j_3 3-arcs. We set $G_k(n,\ell,j_1,j_2,j_3) = |\mathcal{G}_{n,k}(\ell,j_1,j_2,j_3)|$. In particular, we have $G_k(n,\ell,0,0,0) = \mathsf{T}_k^{[4]}(n,\ell)$. We observe that Claim 1 is implied (taking the sum over all ℓ) by

$$\mathsf{T}_k^{[4]}(n,\ell) = \sum_{b\leq\lfloor\frac{n}{2}\rfloor}(-1)^b\lambda(n,b)\,f_k(n-2b,\ell), \tag{5.1}$$

where $\lambda(n,b)$ satisfies the recursion

$$\begin{aligned}\lambda(n,b) &= \lambda(n-1,b) + \lambda(n-4,b-2) + \lambda(n-5,b-2) + \lambda(n-6,b-3)\\ &\quad + \sum_{i=1}^{b}[\lambda(n-2i,b-i) + 2\lambda(n-2i-1,b-i) + \lambda(n-2i-2,b-i)]\\ &\quad - \lambda(n-3,b-1)\end{aligned} \tag{5.2}$$

with the initial conditions $\lambda(n,0) = 1$, $\lambda(n,1) = 3n-6$ and $n \geq 2b$.

We shall proceed by proving Equation (5.1). For this purpose, let $\lambda(n,b_1,b_2,b_3)$ denote the number of ways to select exactly b_1 1-arcs, b_2 2-arcs, and b_3 3-arcs over $1,\ldots,n$ vertices.

Claim A.

$$\sum_{j_1\geq b_1, j_2\geq b_2, j_3\geq b_3} \binom{j_1}{b_1}\binom{j_2}{b_2}\binom{j_3}{b_3} G_k(n,\ell,j_1,j_2,j_3) = \lambda(n,b_1,b_1,b_3)\, f_k(n-2(b_1+b_2+b_3),\ell). \quad (5.3)$$

The idea is to construct a family $\mathcal{F}$ of $\mathcal{G}_{n,k}$-diagrams, having ℓ isolated points and at least b_1 1-arcs, b_2 2-arcs, and b_3 3-arcs, respectively. We then express $|\mathcal{F}|$ via the numbers $G_k(n,\ell,j_1,j_2,j_3)$. We select (a) b_1 1-arcs and b_2 2-arcs and b_3 3-arcs and (b) an arbitrary k-noncrossing diagram over the remaining $n-2(b_1+b_2+b_3)$ vertices with exactly ℓ isolated points. Let $\mathcal{F}$ be the family of diagrams obtained in this way. It is straightforward to show that $\lambda(n,b_1,b_2,b_3)$ satisfies the recursion:

$$\begin{aligned}\lambda(n,b_1,b_2,b_3) = {} & \lambda(n-1,b_1,b_2,b_3) + \lambda(n-2,b_1-1,b_2,b_3) + \lambda(n-4,b_1-1,b_2,b_3-1)\\ & + \lambda(n-5,b_1,b_2,b_3-2) + \lambda(n-6,b_1,b_2,b_3-3) - \lambda(n-3,b_1,b_2-1,b_3)\\ & + \sum_{i=1}^{b}[2\lambda(n-2i-1,b_1,b_2-1,b_3-(i-1)) + \lambda(n-2i-2,b_1,b_2,b_3-i)]\\ & + \sum_{i=2}^{b}[\lambda(n-2i,b_1,b_2-2,b_3-(i-2))]\end{aligned}$$

with the initial conditions $\lambda(n,0,0,0)=1$, $\lambda(n,1,0,0)=n-1$, $\lambda(n,0,1,0)=n-2$, $\lambda(n,0,0,1)=n-3$, $n\geq 2b$.

Clearly, each element $\theta\in\mathcal{F}$ is contained in $\mathcal{G}_{n,k}(\ell,j_1,j_2,j_3)$ for some $j_1\geq b_1$ and $j_2\geq b_2$ and $j_3\geq b_3$. Indeed, any 1-arc or 2-arc or 3-arc can only cross at most two other arcs. Therefore, 1-arcs and 2-arcs and 3-arcs cannot be contained in a set of more than 3-mutually crossing arcs. As a result, for $k>3$ the construction generates k-noncrossing diagrams. Clearly, θ has exactly ℓ isolated vertices and in step (b) we potentially derive additional 1-arcs and 2-arcs and 3-arcs, whence $j_1\geq b_1$ and $j_2\geq b_2$ and $j_3\geq b_3$, respectively. Next we observe that we have by construction

$$|\mathcal{F}| = \lambda(n,b_1,b_2,b_3)\, f_k(n-2(b_1+b_2+b_3),\ell).$$

Since any of the k-noncrossing diagrams over $n-2(b_1+b_2+b_3)$ vertices can generate additional 1-arcs or 2-arcs or 3-arcs, we consider

$$\mathcal{F}(j_1,j_2,j_3) = \{\theta\in\mathcal{F} \mid \theta \text{ has exactly } j_1 \text{ 1-arcs, } j_2 \text{ 2-arcs and } j_3 \text{ 3-arcs}\}.$$

Obviously, we then have the partition $\mathcal{F} = \dot\bigcup_{j_1\geq b_1, j_2\geq b_2, j_3\geq b_3}\mathcal{F}(j_1,j_2,j_3)$. Suppose $\theta\in\mathcal{F}(j_1,j_2,j_3)$, then $\theta\in\mathcal{G}_{n,k}(\ell,j_1,j_2,j_3)$ and furthermore θ occurs with multiplicity $\binom{j_1}{b_1}\binom{j_2}{b_2}\binom{j_3}{b_3}$ in $\mathcal{F}$ since by construction any b_1-element subset of the j_1 1-arcs and b_2-element subset of the j_2 2-arcs and b_3-element subset of the j_3 3-arcs is counted respectively in $\mathcal{F}$. Therefore, we have

$$|\mathcal{F}(j_1,j_2,j_3)| = \binom{j_1}{b_1}\binom{j_2}{b_2}\binom{j_3}{b_3} G_k(n,\ell,j_1,j_2,j_3) \quad (5.4)$$

and

$$\sum_{j_1\geq b_1, j_2\geq b_2, j_3\geq b_3} |\mathcal{F}(j_1,j_2,j_3)| = \lambda(n,b_1,b_2,b_3)\, f_k(n-2(b_1+b_2+b_3),\ell)$$

proving Claim A. We next set

$$F_k(x,y,z) = \sum_{j_1\geq 0}\sum_{j_2\geq 0}\sum_{j_3\geq 0} G_k(n,\ell,j_1,j_2,j_3)x^{j_1}y^{j_2}z^{j_3}.$$

Taking derivatives, we obtain

$$\frac{1}{b_1!}\frac{1}{b_2!}\frac{1}{b_3!}F_k^{(b_1,b_2,b_3)}(1) = \sum_{j_1\geq b_1, j_2\geq b_2, j_3\geq b_3} \binom{j_1}{b_1}\binom{j_2}{b_2}\binom{j_3}{b_3} G_k(n,\ell,j_1,j_2,j_3)1^{j_1-b_1}1^{j_2-b_2}1^{j_3-b_3}$$

and accordingly

$$\sum_{j_1\geq 0, j_2\geq 0, j_3\geq 0} G_k(n,\ell,j_1,j_2,j_3)x^{j_1}y^{j_2}z^{j_3}$$

$$= \sum_{b_1\geq 0, b_2\geq 0, b_3\geq 0}\left[\sum_{j_1\geq b_1, j_2\geq b_2, j_3\geq b_3} \binom{j_1}{b_1}\binom{j_2}{b_2}\binom{j_3}{b_3} G_k(n,\ell,j_1,j_2,j_3)\right]$$

$$(x-1)^{b_1}(y-1)^{b_2}(z-1)^{b_3}$$

$$= \sum_{b_1\geq 0, b_2\geq 0, b_3\geq 0} \lambda(n,b_1,b_2,b_3)\, f_k(n-2(b_1+b_2+b_3),\ell)(x-1)^{b_1}(y-1)^{b_2}(z-1)^{b_3}.$$

By construction, $G(n,\ell,0,0,0)$ is the constant term of the $F_k(x,y,z)$. That is, the number of k-noncrossing RNA structures with ℓ isolated vertices and no 1-arcs, 2-arcs, and 3-arcs is given by

$$G(n,\ell,0,0,0) = \sum_{b_1\geq 0, b_2\geq 0, b_3\geq 0} (-1)^{b_1+b_2+b_3}\lambda(n,b_1,b_2,b_3)\, f_k(n-2(b_1+b_2+b_3),\ell). \tag{5.5}$$

We take the sum over all ℓ and derive

$$\mathsf{T}_k^{[4]}(n) = \sum_{b_1\geq 0, b_2\geq 0, b_3\geq 0}^{\lfloor \frac{n}{2}\rfloor} (-1)^{b_1+b_2+b_3}\lambda(n,b_1,b_2,b_3)\left[\sum_{\ell=0}^{n-2(b_1+b_2+b_3)} f_k(n-2(b_1+b_2+b_3),\ell)\right]. \tag{5.6}$$

Setting

$$\lambda(n,b) = \sum_{b_1+b_2+b_3=b} \lambda(n,b_1,b_2,b_3)$$

we conclude first

$$\mathsf{T}_k^{[4]}(n) = \sum_{b\leq \lfloor \frac{n}{2}\rfloor} (-1)^b \lambda(n,b)\,\mathsf{M}_k(n-2b)$$

and second Equation (5.2), completing the proof of Claim 1.

ACKNOWLEDGMENTS

We are grateful to Fenix W.D. Huang, Emma Y. Jin, Jing Qin, and Rita R. Wang for their help. This work was supported by the 973 Project, the PCSIRT Project of the Ministry of Education, the Ministry of Science and Technology, and the National Science Foundation of China.

DISCLOSURE STATEMENT

No competing financial interests exist.

REFERENCES

Abramowitz, M., and Stegun, I.A., eds. 1964. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. NBS Applied Mathematics Series 55. National Bureau of Standards, Washington, D.C.

Chen, W.Y.C., Deng, E.Y.P., Du, R.R.X., et al. 2007. Crossings and nestings of matchings and partitions. *Trans. Am. Math. Soc.* 359, 1555–1575.

Chen, W.Y.C., Qin, J., and Reidys, C.M. 2008a. Crossing and nesting of tangled-diagrams. *Electron. J. Combin.* 15, 86.

Chen, W.Y.C., Qin, J., Reidys, C.M., et al. 2008b. Efficient counting and asymptotics of k-noncrossing tangled diagrams (submitted).

Egorychev, G.P. 1984. *Integral Representation and the Computation of Combinatorial Sums. Translations of Mathematical Monographs, Vol. 59*. American Mathematical Society, Providence, RI.

Flajolet, P., and Sedgewick, R. 2008. *Analytic Combinatorics*. Cambridge University Press, Cambridge, UK.

Gessel, I., and Zeilberger, D. 1992. Random walk in a Weyl chamber. *Proc. Am. Math. Soc.* 115, 27–31.

Grabiner, D.J., and Magyar, P. 1993. Random walks in Weyl chambers and the decomposition of tensor powers. *Discrete Appl. Math.* 2, 239–260.

Haslinger, C., and Stadler, P.F. 1999. RNA structures with pseudo-knots. *Bull. Math. Biol.* 61, 437–467.

Hofacker, I.L., Schuster, P., and Stadler, P.F. 1998. Combinatorics of RNA secondary structures. *Discrete Appl. Math.* 88, 207–237.

Howell, J.A., Smith, T.F., and Waterman, M.S. 1980. Computation of generating functions for biological molecules. *SIAM J. Appl. Math.* 39, 119–133.

Jin, E.Y., Qin, J., and Reidys, C.M. 2008a. Combinatorics of RNA structures with pseudoknots. *Bull. Math. Biol.* 70, 45–67.

Jin, E.Y., and Reidys, C.M. 2008a. Asymptotic enumeration of RNA structures with pseudoknots. *Bull. Math. Biol.* 70, 951–970.

Jin, E.Y., and Reidys, C.M. 2008b. Combinatorial design of pseudoknot RNA. *Adv. Appl. Math.* (in press).

Jin, E.Y., Reidys, C.M., and Wang, R.R. 2008b. Asympotic analysis of k-noncrossing matchings. Submitted.

Konings, D.A.M., and Gutell, R.R. 1995. A comparison of thermodynamic foldings with comparatively derived structures of 16S and 16S-like rRNAs. *RNA* 1, 559–574.

Loria, A., and Pan, T. 1996. Domain structure of the ribozyme from eubacterial ribonuclease P. *RNA* 2, 551–563.

Mapping RNA. 2005. Mapping RNA form and function. *Science* [Special Issue] 2.

Odlyzko, A.M. 1995. Asymptotic enumeration methods, 1063–1229. In: Graham, R.L., Groetschel, M., and Lovasz, L., eds. *Handbook of Combinatorics. Vol. 2*. Elsevier, New York.

Penner, R.C., and Waterman, M.S. 1993. Spaces of RNA secondary structures. *Adv. Math.* 101, 31–49.

Rivas, E., and Eddy, S. 1999 A dynamic programming algorithm for RNA structure prediction inclusing pseudoknots. *J. Mol. Biol.* 285, 2053–2068.

Stanley, R. 1980. Differentiably finite power series. *Eur. J. Combinatorics* 1, 175–188.

Titchmarsh, E.C. 1939. *The Theory of Functions*. Oxford University Press, Oxford, UK.

Wasow, W. 1987. *Asymptotic Expansions for Ordinary Differential Equations*. Dover, New York.

Waterman, M.S. 1978. Secondary structure of single-stranded nucleic acids. *Adv. Math. Suppl. Stud. I* 1, 167–212.

Waterman, M.S. 1979. Combinatorics of RNA hairpins and cloverleafs. *Stud. Appl. Math.* 60, 91–96.

Waterman, M.S., and Schmitt, W.R. 1994. Linear trees and RNA secondary structure. *Discrete Appl. Math.* 51, 317–323.

Westhof, E., and Jaeger, L. 1992. RNA pseudoknots. *Curr. Opin. Struct. Biol.* 2, 327–333.

Wong, R., and Wyman, M. 1974. The method of Darboux. *J. Approx. Theory* 10, 159–171.

Zeilberger, D. 1990. A holonomic systems approach to special functions identities. *J. Comput. Appl. Math.* 32, 321–368.

Address reprint requests to:
Dr. Christian M. Reidys
Center for Combinatorics, LPMC-TJKLC
Nankai University
Tianjin 300071, P.R. China

E-mail: reidys@nankai.edu.cn