Closed-form expressions for the moments of the binomial probability distribution.

Andreas Knoblauch

2008

Preprint:

This is an accepted article published in SIAM Journal on Applied Mathematics. The final authenticated version is available online at: https://doi.org/[DOI not available]

CLOSED-FORM EXPRESSIONS FOR THE MOMENTS OF THE BINOMIAL PROBABILITY DISTRIBUTION *

ANDREAS KNOBLAUCH†

Abstract. This work develops closed form expressions for the raw and central moments of the binomial probability distribution. For this I first derive a recursive formula for the raw moments from the moment generating function. Then it is shown that the recursion involved is essentially the same as for the Stirling numbers of the second kind. From this fact it is then possible to derive the closed formulae. Finally, I discuss an application of these formulae to the analysis of neural associative memory.

Key words. Stirling numbers, Pochhammer polynomial, neural associative memory, storage capacity, retrieval errors

AMS subject classifications. 05A15, 60C05, 92B20

1. The binomial probability and its moments. A random variable X is called *binomially distributed* with parameters n and p if the random variable takes value $x \in \{0, 1, 2, \ldots, n\}$ with probability

$$p_B(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x} \ . \tag{1.1}$$

The *moment generating function* $G_B(s) := E_{p_B} e^{sX}$ of the binomial probability can then be computed using the binomial sum $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$,

$$G_B(s; n, p) = \sum_{x=0}^{n} \binom{n}{x} (pe^s)^x (1-p)^{n-x} = (pe^s + 1 - p)^n \ . \tag{1.2}$$

The d-th *raw moment* $E_{p_B} X^d$ equals the d-th derivative of the generating function $G_B(s)$ at $s = 0$ (e.g., [14]). For example, the mean value is $\mu := E_{p_B} X = n(pe^s + 1-p)^{n-1} pe^s|_{s=0} = np$ and the second raw moment is $E_{p_B} X^2 = np((n-1)(pe^s + 1-p)^{n-2} pe^s + (pe^s + 1-p)^{n-1} e^s)|_{s=0} = np(np + 1 - p)$. Higher-order moments for larger d can be computed, in principle, by continuing this procedure, but computing higher-order derivatives of $G_B(s)$ becomes tedious with increasing d. In the following we aim at finding a recursive formula without referring to higher order derivatives of $G_B(s)$ (see also [2] for a related approach).

2. Recursive formulae. For computing the higher-order derivatives of the moment generating function $G_B(s)$ for larger d we can define auxiliary functions

$$H_d(s) := (pe^s)^d \text{ with derivative } H'_d(s) = dH_d(s) \ , \tag{2.1}$$

$$F_d(s) := n^{\underline{d}} G_B(s; n-d, p) \text{ with derivative} \tag{2.2}$$

$$F'_d(s) = n^{\underline{d+1}} (pe^s + 1 - p)^{n-d-1} H_1(s) = F_{d+1}(s) H_1(s) \ , \text{ and} \tag{2.3}$$

$$K_d(s) := F_d(s) H_d(s) \text{ with derivative} \tag{2.4}$$

$$K'_d(s) = F'_d(s) H_d(s) + dF_d(s) H_d(s) = K_{d+1}(s) + dK_d(s) \ , \tag{2.5}$$

*Published in: SIAM Journal on Applied Mathematics 69(1), pp 197–204, 2008.

†Honda Research Institute Europe, Carl-Legien-Strasse 30, D-63073 Offenbach/Main, Germany (andreas.knoblauch@honda-ri.de).

where $n^{\underline{d}} = n(n-1)\cdots(n-d+1)$ denotes a falling factorial or Pochhammer symbol. Since $G_B(s) = K_0(s)$ we can obtain the higher-order derivatives of the moment generating function $G_B(s)$ recursively from eq. 2.5, for example $G_B^{(0)} = K_0$, $G_B^{(1)} = K_1$, $G_B^{(2)} = K_2 + K_1$, $G_B^{(3)} = K_3 + 2K_2 + K_2 + K_1 = K_3 + 3K_2 + K_1$. Thus, we can prove the following

LEMMA 2.1. *The d-th derivative $G_B^{(d)}$ of the moment generating function $G_B(s)$ of the binomial probability $p_B(x; n, p)$ can be written as a weighted sum of functions $K_i(s)$,*

$$G_B^{(d)}(s) = \sum_{i=0}^{d} b_{di} K_i(s) \ . \tag{2.6}$$

for appropriate coefficients b_{di}. The coefficients can be computed recursively from

$$b_{0i} = \delta_{i0} \tag{2.7}$$

$$b_{di} = i b_{d-1,i} + b_{d-1,i-1} \ . \tag{2.8}$$

where δ_{ij} is the usual Kronecker symbol (1 for $i = j$, and 0 otherwise). For convenience we further define $b_{di} = 0$ for $d < 0$ or $i < 0$ or $i > d$.

Proof. Eq. 2.7 follows from $G_B^{(0)} = K_0$ (see eq. 2.4). Eq. 2.8 can then be shown inductively using eq. 2.6 with eq. 2.5,

$$\begin{aligned} G_B^{(d+1)}(s) &= \sum_{i=0}^{d} b_{di}(K_{i+1} + iK_i) = \sum_{i=1}^{d+1} b_{d,i-1} K_i + \sum_{i=0}^{d} b_{di} i K_i \\ &= \sum_{i=0}^{d+1} (i b_{di} + b_{d,i-1}) K_i \ . \end{aligned}$$

□

From this lemma and $K_i(0) = n^{\underline{i}} p^i$ we can give recursive formulae for the raw and central moments as summarized by the following theorem.

THEOREM 2.2. *The d-th raw and central moment of a binomially distributed random variable X with $\mathrm{pr}[X = x] = p_B(x; n, p)$, expectation $\mu := np$, and $q := 1 - p$ are*

$$E_{p_B} X^d = \sum_{i=0}^{d} b_{di} p^i n^{\underline{i}} \ . \tag{2.9}$$

$$= \sum_{j=0}^{d} (-q)^j \sum_{i=j}^{d} \binom{i}{j} b_{di} n^{\underline{i}} \tag{2.10}$$

$$E_{p_B}(X - \mu)^d = \sum_{i=0}^{d} \binom{d}{i} (-\mu)^{d-i} E_{p_B} X^i \ . \tag{2.11}$$

Eqs. 2.10,2.11 can be obtained from the binomial sum (see section 1). Eq. 2.10 is written as polynomial in q which is useful for some applications (see section 5). The first few values of the coefficients b_{di} are shown in table 2.1.

b_{di}	i=0	1	2	3	4	5	6	7	8	9	10
d=0	1	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0	0	0	0
3	0	1	3	1	0	0	0	0	0	0	0
4	0	1	7	6	1	0	0	0	0	0	0
5	0	1	15	25	10	1	0	0	0	0	0
6	0	1	31	90	65	15	1	0	0	0	0
7	0	1	63	301	350	140	21	1	0	0	0
8	0	1	127	966	1701	1050	266	28	1	0	0
9	0	1	255	3025	7770	6951	2646	462	36	1	0
10	0	1	511	9330	34105	42525	22827	5880	750	45	1

TABLE 2.1

Values of the binomial moment coefficients b_{di} for $0 \le d, i \le 10$. These coefficients can be used to compute the moments of the binomial probability (see eq. 2.9), and are identical to the Stirling numbers of the second kind (see section 3).

3. Relation to Stirling numbers of the second kind. The coefficients b_{di} for computing the binomial moments (eq. 2.9) are actually Stirling numbers of the second kind: The Stirling number of the second kind $S(d,i)$ is defined as the number of ways of partitioning a set of d elements into i nonempty sets, and one can show that $S(d,i)$ obeys the same recurrence relations eqs. 2.7,2.8 as b_{di} (e.g., [1, 17, 9]). Closed formula for the Stirling numbers of the second kind are well known, for example,

$$b_{di} = S(d,i) = \frac{(-1)^i}{i!} \sum_{k=0}^{i} (-1)^k \binom{i}{k} k^d \,. \tag{3.1}$$

Thus, inserting this into the formulae of theorem 2.2 gives us already closed-form expressions for the moments of the binomial probability. However, these formulae can still be simplified using a generalization of the following generating function (e.g., see [17])

$$n^d = \sum_{i=0}^{d} b_{di} n^{\underline{i}} \,. \tag{3.2}$$

The generalization is given by the following lemma.

LEMMA 3.1.

$$\sum_{i=j}^{d} \binom{i}{j} b_{di} n^{\underline{i}} = \binom{n}{j} \sum_{k=0}^{j} (-1)^k \binom{j}{k} (n-k)^d \,. \tag{3.3}$$

Proof. Instead of eq. 3.3 we prove the equivalent equation

$$C(d,j,n) := \sum_{i=j}^{d} i^{\underline{j}} b_{di} n^{\underline{i}} = n^{\underline{j}} \sum_{k=0}^{j} (-1)^k \binom{j}{k} (n-k)^d \tag{3.4}$$

by induction over j. For $j = 0$ the lemma is identical to eq. 3.2. For larger j we compute using $ib_{di} = b_{d+1,i} - b_{d,i-1}$ (see eq. 2.8)

$$C(d,j+1,n) = \sum_{i=j+1}^{d} i^{\underline{j+1}} b_{di} n^{\underline{i}} = \sum_{i=j+1}^{d} i^{\underline{j}} i b_{di} n^{\underline{i}} - j \sum_{i=j+1}^{d} i^{\underline{j}} b_{di} n^{\underline{i}}$$

$$= \sum_{i=j+1}^{d} i^{\underline{j}} b_{d+1,i} n^{\underline{i}} - \sum_{i=j+1}^{d} i^{\underline{j}} b_{d,i-1} n^{\underline{i}} - j \sum_{i=j+1}^{d} i^{\underline{j}} b_{di} n^{\underline{i}} .$$

The three sums can be written individually

$$\begin{aligned}
S_1 &:= \sum_{i=j+1}^{d} i^{\underline{j}} b_{d+1,i} n^{\underline{i}} = C(d+1,j,n) - j! b_{d+1,j} n^{\underline{j}} - (d+1)^{\underline{j}} n^{\underline{d+1}} \\
S_2 &:= \sum_{i=j+1}^{d} i^{\underline{j}} b_{d,i-1} n^{\underline{i}} \\
&= n \sum_{i=j+1}^{d} (i-1)^{\underline{j}} b_{d,i-1} (n-1)^{\underline{i-1}} + jn \sum_{i=j+1}^{d} (i-1)^{\underline{j-1}} b_{d,i-1} (n-1)^{\underline{i-1}} \\
&= n \sum_{i=j}^{d-1} i^{\underline{j}} b_{d,i} (n-1)^{\underline{i}} + jn \sum_{i=j}^{d-1} i^{\underline{j-1}} b_{d,i} (n-1)^{\underline{i}} \\
&= nC(d,j,n-1) - nd^{\underline{j}} (n-1)^{\underline{d}} \\
&\quad + jnC(d,j-1,n-1) - jn(j-1)^{\underline{j-1}} b_{d,j-1} (n-1)^{\underline{j-1}} - jnd^{\underline{j-1}} (n-1)^{\underline{d}} \\
S_3 &:= j \sum_{i=j+1}^{d} i^{\underline{j}} b_{di} n^{\underline{i}} = jC(d,j,n) - j j^{\underline{j}} b_{dj} n^{\underline{j}} ,
\end{aligned}$$

where we used $b_{dd} = 1$ for $d \geq 0$. For the second sum S_2 we used $i^{\underline{j}} = ((i-j)+j)(i-1)^{\underline{j-1}}$. Fortunately, in $S_1 - S_2 - S_3$ all the non-C terms cancel out: The b terms cancel out because with $b_{d+1,j} = jb_{d,j} + b_{d,j-1}$ (eq. 2.8) we have

$$-j! n^{\underline{j}} b_{d+1,j} + j! n^{\underline{j}} b_{d,j-1} + jj! n^{\underline{j}} b_{dj} = j! n^{\underline{j}} (-b_{d+1,j} + b_{d,j-1} + jb_{dj}) = 0 .$$

The remaining non-C and non-b terms cancel out because

$$-(d+1)^{\underline{j}} n^{\underline{d+1}} + d^{\underline{j}} n^{\underline{d+1}} + jd^{\underline{j-1}} n^{\underline{d+1}} = n^{\underline{d+1}} d^{\underline{j-1}} (-(d+1) + (d-j+1) + j) = 0 .$$

Thus, using the induction hypothesis, we have simply

$$\begin{aligned}
C(d,j+1,n) &= C(d+1,j,n) - nC(d,j,n-1) - jnC(d,j-1,n-1) - jC(d,j,n) \\
&= n^{\underline{j}} \sum_{k=0}^{j} \binom{j}{k} (-1)^k (n-k)^{d+1} \\
&\quad - n(n-1)^{\underline{j}} \sum_{k=0}^{j} \binom{j}{k} (-1)^k (n-(k+1))^d \\
&\quad - jn(n-1)^{\underline{j-1}} \sum_{k=0}^{j-1} \binom{j-1}{k} (-1)^k (n-(k+1))^d \\
&\quad - jn^{\underline{j}} \sum_{k=0}^{j} \binom{j}{k} (-1)^k (n-k)^d \\
&= \sum_{k=0}^{j+1} a_k (n-k)^d .
\end{aligned}$$

In the last line we have simply summed over the $(n-k)^d$ terms. Thus, our proof is finished if we can show that $a_k = n^{\underline{j+1}}\binom{j+1}{k}(-1)^k$ for $k = 0, 1, \ldots, j+1$. The highest coefficient a_{j+1} gets contributions only from the second sum,

$$a_{j+1} = -n^{\underline{j+1}}(-1)^j(n-(j+1))^d = n^{\underline{j+1}}\binom{j+1}{j+1}(-1)^k \ .$$

The lowest coefficient a_0 gets contributions only from the first and fourth sum,

$$a_0 = n^{\underline{j}}n - jn^{\underline{j}} = n^{\underline{j+1}} = n^{\underline{j+1}}\binom{j+1}{0}(-1)^0 \ .$$

The remaining intermediary coefficients a_k for $k = 1, 2, \ldots, j$ get contributions from all four sums,

$$\begin{aligned}
a_k &= n^{\underline{j}}\binom{j}{k}(-1)^k(n-k) - n^{\underline{j+1}}\binom{j}{k-1}(-1)^{k-1} \\
&\quad -jn^{\underline{j}}\binom{j-1}{k-1}(-1)^{k-1} - jn^{\underline{j}}\binom{j}{k}(-1)^k \\
&= (-1)^k n^{\underline{j}}\binom{j}{k}((n-k) + (n-j)\frac{k}{j-k+1} + j\frac{k}{j} - j) \\
&= (-1)^k n^{\underline{j}}\binom{j}{k}\frac{(n-j)(j+1)}{j-k+1} = n^{\underline{j+1}}\binom{j+1}{k}(-1)^k \ .
\end{aligned}$$

Thus, we have proven eq. 3.4. □

A useful variant of lemma 3.1 including an offset μ is

$$\binom{n}{j}\sum_{k=0}^{j}(-1)^k\binom{j}{k}(n-\mu-k)^d = n^{\underline{j}}\sum_{i=j}^{d}\binom{i}{j}b_{di}(n-\mu-j)^{\underline{i-j}} \ . \tag{3.5}$$

4. Closed formulae. The following theorem summarizes the main results of this work:

THEOREM 4.1. *Let X be a binomially distributed random variable with probability function $p_B(x; n, p)$ (see eq. 1.1). Further let $q := 1-p$ and let b_{di} Stirling numbers of the second kind (see table 2.1 and eqs. 2.7,2.8,3.1). Then the d-th raw moment of the binomial probability p_B can be written*

$$E_{p_B}X^d = \sum_{i=0}^{d} b_{di}p^i n^{\underline{i}} \tag{4.1}$$

$$= \sum_{i=0}^{d}(-p)^i\binom{n}{i}\sum_{k=0}^{i}(-1)^k\binom{i}{k}k^d \tag{4.2}$$

$$= \sum_{j=0}^{d}(-q)^j\sum_{i=j}^{d}\binom{i}{j}b_{di}n^{\underline{i}} \tag{4.3}$$

$$= \sum_{j=0}^{d}(-q)^j\binom{n}{j}\sum_{k=0}^{j}(-1)^k\binom{j}{k}(n-k)^d \ . \tag{4.4}$$

For the d-th raw moment the following bounds are true,

$$(np)^d \leq E_{p_B}X^d \leq n^d \ . \tag{4.5}$$

For an arbitrary offset μ we have

$$E_{p_B}(X-\mu)^d = \sum_{i=0}^{d} \binom{d}{i} (-\mu)^{d-i} E_{p_B} X^i \tag{4.6}$$

$$= \sum_{j=0}^{d} (-p)^j \binom{n}{j} \sum_{k=0}^{j} (-1)^k \binom{j}{k} (k-\mu)^d \tag{4.7}$$

$$= \sum_{j=0}^{d} (-q)^j n^{\underline{j}} \sum_{i=j}^{d} \binom{i}{j} b_{di} (n-\mu-j)^{\underline{i-j}} \tag{4.8}$$

$$= \sum_{j=0}^{d} (-q)^j \binom{n}{j} \sum_{k=0}^{j} (-1)^k \binom{j}{k} (n-\mu-k)^d \ . \tag{4.9}$$

In particular, for $\mu = E_{p_B} X = np$ we obtain the d-th central moment of the binomial probability p_B. For $z - d \geq \mu \geq E_{p_B} X$ the following bound is true,

$$|E_{p_B}(X-\mu)^d| \leq \sum_{j=0}^{d} \sum_{i=j}^{d} \binom{i}{j} b_{di} (nq)^i \quad (\sim (nq)^d \text{ for fixed } d \text{ and } nq \to \infty) \ . \tag{4.10}$$

Proof. Eqs. 4.1,4.2 follow from eqs. 2.9,3.1. Eqs. 4.3,4.4 follow from eqs. 2.10,3.3. The bounds of eq. 4.5 follows simply from $p^d \leq p^i \leq 1$ and eq. 3.2 because eq. 4.1 is obviously a sum of non-negative numbers. Eq. 4.6 is eq. 2.11. Eq. 4.7 follows by inserting eq. 4.2 into 4.6 and applying the binomial sum (see section 1). Similarly, eq. 4.9 follows from inserting eq. 4.4 into eq. 4.6. Eq. 4.8 follows from eq. 4.9 with eq 3.5. Eq. 4.10 follows from eq. 4.8 because $n^{\underline{j}} \leq n^j$ and $(n-\mu-j)^{\underline{i-j}} \leq (nq)^{i-j}$ for $n \geq \mu + d$ and $\mu \geq E_{p_B} X = np$. □

5. Related work and application to the analysis of neural associative networks. Computing the higher-order moments of a binomially distributed random variable is rarely emphasized. Standard textbooks give expressions for the moment generating function (eq. 1.2) and some lower-order moments such as mean, variance, and, perhaps, skewness and kurtosis, but higher-order moments are usually neglected (e.g., see [14, 16]). For some applications it may be sufficient to approximate a binomial random variable by either a Gaussian or a Poissonian where closed-form expressions for higher-order moments are known. For example, for large variance $np(1-p) \to \infty$, according to the DeMoivre-Laplace theorem, the binomial probability becomes similar to a Gaussian with same mean and variance. Likewise, for $n \to \infty$ and finite $np \to \lambda < \infty$ the binomial becomes Poissonian. However, for applications as described below, these approximations are not appropriate and it is necessary to find an exact formula.

A previous attempt [2] to compute the higher-order moments of the binomial distribution revealed recursive expressions similar to those developed in section 2, but was restricted to the special case $p = 0.5$. Moreover, the recursive form was not appropriate for efficient computation or application in further analyses. In contrast to [2], this work provides general recursive and non-recursive (or closed-form) expressions for the higher-order moments of the binomial distribution.

My main motivation to obtain a closed formula for the binomial moments comes from analyzing storage capacity and retrieval error probabilities in neural associative memory networks [21, 12, 4, 18, 6]. *Associative memories* are systems that contain information about a finite set of associations between pattern vector pairs $\{(\mathbf{u}^\mu \mapsto \mathbf{v}^\mu) : \mu = 1, ..., M\}$ [10]. Given a possibly noisy address pattern $\tilde{\mathbf{u}}$ the problem is to find a target pattern $\mathbf{v}^\mu$ for which the corresponding address pattern $\mathbf{u}^\mu$ is most similar to $\tilde{\mathbf{u}}$.

Neural associative networks have wide applications both for artificial intelligence (e.g., visual object recognition [10, 15]) and modeling the brain (e.g., [13, 6, 22, 5, 19, 20]). In neural implementations the associations are stored in a matrix $\mathbf{A}$ describing the synaptic connections between two cell populations u and v. Here the retrieval result $\hat{\mathbf{v}}^\mu$ may differ from the original pattern $\mathbf{v}^\mu$. This is due to retrieval noise being an increasing function of the memory load or the number of stored associations. In general the probability of a retrieval error can be computed from the neuron potential distribution as obtained by propagating the address pattern $\tilde{\mathbf{u}}$ through the synaptic matrix $\mathbf{A}$.

One of the most efficient models is the so-called *Willshaw network* with binary neurons and synapses [21]. Here the synaptic matrix is simply $\mathbf{A} = \vee_{\mu=1}^{M} \mathbf{u}^{\mu,\mathbf{T}} \mathbf{v}^\mu$ and the retrieval error probabilities can be computed from the so-called Willshaw-Palm distribution of neuron potentials $\mathbf{x} = \tilde{\mathbf{u}}^{\mathbf{T}} \mathbf{A}$. Since the Willshaw-Palm distribution is more difficult to formulate, many analyses of neural associative memory actually rely on a binomial approximation (e.g., [21, 12, 11, 3, 18]). However, it is unclear for which network parameters this approximation is sufficiently accurate. In a further paper [8] (see also [7]) I will compute the moments of the Willshaw-Palm distribution from the binomial moments. For this it is sufficient to replace q^j in eqs. 4.4,4.9 by some more complex term $q^{(j)}$. With this it will be possible to compare the exact potential distribution to the binomial approximation and determine asymptotic conditions when they become identical.

Acknowledgments. I am grateful to Edgar Körner and Marc-Oliver Gewaltig for providing the opportunity to do this work at the Honda Research Institute. I am also grateful to Günther Palm and Friedrich Sommer for helpful discussions and comments.

REFERENCES

[1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing.* Dover, New York, 1972.

[2] A. Benyi and S.M. Manago. A recursive formula for moments of a binomial distribution. *College Mathematics Journal*, 36(1):68–72, 2005.

[3] H. Bosch and F. Kurfess. Information storage capacity of incompletely connected associative memories. *Neural Networks*, 11(5):869–876, 1998.

[4] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Science, USA*, 79:2554–2558, 1982.

[5] R.A. Jortner, S.S. Farivar, and G. Laurent. A simple connectivity scheme for sparse coding in an olfactory system. *The Journal of Neuroscience*, 27(7):1659–1669, 2007.

[6] A. Knoblauch. Neural associative memory for brain modeling and information retrieval. *Information Processing Letters*, 95:537–544, 2005.

[7] A. Knoblauch. Asymptotic conditions for high-capacity neural associative networks. Internal Report HRI-EU 07-02, Honda Research Institute Europe GmbH, D-63073 Offenbach/Main, Germany, February 2007.

[8] A. Knoblauch. Neural associative memory and the Willshaw-Palm probability distribution. *SIAM Journal on Applied Mathematics*, 69(1):169–196, 2008.

[9] D.E. Knuth. *The Art of Computer Programming, Volume 1: Fundamental Algorithms, 3rd Edition.* Addison-Wesley, Reading, MA, 1997.

[10] T. Kohonen. *Associative memory: a system theoretic approach.* Springer, Berlin, 1977.

[11] J.-P. Nadal. Associative memory: on the (puzzling) sparse coding limit. *J.Phys. A: Math. Gen.*, 24:1093–1101, 1991.
[12] G. Palm. On associative memories. *Biological Cybernetics*, 36:19–31, 1980.
[13] G. Palm. *Neural Assemblies. An Alternative Approach to Artificial Intelligence.* Springer, Berlin, 1982.
[14] A. Papoulis. *Probability, Random Variables, and Stochastic Processes. Third edition.* McGraw-Hill, New York, 1991.
[15] M. Rehn and F.T. Sommer. Storing and restoring visual input with collaborative rank coding and associative memory. *Neurocomputing*, 69:1219–1223, 2006.
[16] J. Riordan. *An introduction to combinatorial analysis.* Wiley, New York, 1958.
[17] S. Roman. *The Umbral Calculus.* Acamdemic Press, New York, 1984.
[18] F.T. Sommer and G. Palm. Improved bidirectional retrieval of sparse patterns stored by Hebbian learning. *Neural Networks*, 12:281–297, 1999.
[19] L.G. Valiant. Memorization and association on a realistic neural model. *Neural Computation*, 17:527–555, 2005.
[20] L.G. Valiant. A quantitative theory of neural computation. *Biological Cybernetics*, 95(3):205–211, 2006.
[21] D.J. Willshaw, O.P. Buneman, and H.C. Longuet-Higgins. Non-holographic associative memory. *Nature*, 222:960–962, 1969.
[22] X. Xie, R.H.R. Hahnloser, and H.S. Seung. Selectively grouping neurons in recurrent networks of lateral inhibition. *Neural Computation*, 14(11):2627–2646, 2002.