Raising the order of multivariate approximation schemes using supplementary derivative data

Dirk Kraaijpoel\textsuperscript{a,}* , Tristan van Leeuwen\textsuperscript{b}

\textsuperscript{a}Climate and Seismology Department, Royal Netherlands Meteorological Institute (KNMI)
\textsuperscript{b}Faculty of Civil Engineering and Geosciences, Delft University of Technology

Abstract

We propose a generic procedure to raise the approximation order of multivariate approximation schemes using supplementary derivative data. The procedure applies to all schemes that reproduce polynomials to a certain degree, including most common types of (quasi-) interpolation and moving least-squares. For an approximation scheme of order $m$ and a dataset that provides $n$ supplementary orders of derivative data, the procedure results in an approximation order of $m+n$. This is achieved using a modification of the Taylor expansion, the \textit{reduced dual Taylor expansion}, that is applied to the data prior to the evaluation of the scheme. The procedure is easy to implement in existing schemes and is expected to be useful immediately in a wide range of applications.

Keywords: approximation, polynomial reproduction, (quasi-) interpolation, moving least-squares, derivatives, reduced dual Taylor expansion

1. Introduction

Functional approximation schemes are omnipresent in the computational sciences. The reader is referred to a number of excellent text books and (review) papers on interpolation and approximation theory [1, 2], spline functions [3], multivariate polynomial interpolation [4, 5], moving least-squares approximation [6], radial basis functions [7], natural neighbor interpolation [8, 9], polygonal finite-element interpolants [10], nonlinear approximation [11], interpolation in signal and image processing [12] and geometric modeling [13].

Usually, the data available to approximation schemes consist of function values at discrete node locations. The choice in schemes that are able to accommodate derivative data is rather limited. In classical interpolation theory incorporation of derivative data is well known as the Hermite interpolation problem. In the univariate case, solution procedures are well established [1]. However, extensions to arbitrary data configurations in multiple dimensions are difficult to construct [4, 5], and probably less attractive for routine application. In non-interpolating approximation schemes, such as quasi-interpolation and moving least-squares, the incorporation of derivative data appears to be largely unexplored.

*To whom correspondence should be addressed.
Email address: dirk.kraaijpoel@knmi.nl (Dirk Kraaijpoel)
In this paper we propose a practical procedure for incorporating supplementary derivative data in existing approximation schemes. The procedure applies generically to all schemes that are based on function values at discrete nodes and that reproduce polynomials to some degree. This includes most common types of interpolation, quasi-interpolation and moving least-squares. The derivative data must be available at the same nodes as the function values. An early version of the procedure, limited to linear interpolation schemes, was published in [14, Ch. 4]. Recently came to our attention an independent line of research [15, 16, 17, 18] that achieves similar results in slightly different contexts. With respect to these papers our procedure is broader in scope because it aims at general applicability in a wider class of approximation schemes. Our presentation is geared toward direct implementation in existing approximation schemes.

The paper is organized as follows. We start by establishing some notation and terminology, including a definition of the dual Taylor expansion in Definition 2.1. Subsequently we motivate our work by showing that this dual Taylor expansion can — in principle — be used to incorporate supplementary derivative data in arbitrary approximation schemes. A simple example, however, reveals that the result is sub-optimal. We then proceed by introducing the central element of our procedure, the reduced dual Taylor expansion in Definition 4.1. It is this expansion that allows us to raise the order of arbitrary approximation schemes as proven in Theorem 5.1. In the subsequent section we provide examples of how the procedure can be applied to scattered data approximation. We finalize with some additional remarks and a conclusion.

2. Notation and terminology

Function space and domain

We seek to approximate multivariate functions \( f : X \rightarrow \mathbb{R} \) with \( X \subseteq \mathbb{R}^d \). Elements of \( X \) are denoted by \( x \equiv (x_1, x_2, \ldots, x_d) \) or anonymous placeholders (.). The class of suitable functions will be denoted by \( V \), such that \( f \in V \).

In the context of this paper it is not necessary to be precise on the continuity properties of \( f \). The function is referred to as an approximant to \( f \).

Multi-index notation

A multi-index \( \kappa \) is defined as a tuple of non-negative integers, \( \kappa \equiv (\kappa_1, \kappa_2, \ldots, \kappa_d) \in \mathbb{N}_0^d \). Multi-indices follow the standard rules for vectorial addition and subtraction and some specific definitions for norm: \( |\kappa| := \sum_{i=1}^d \kappa_i \), factorial: \( \kappa! := \prod_{i=1}^d \kappa_i! \) and binomial coefficient: \( \binom{\kappa}{\lambda} := \prod_{i=1}^d \binom{\kappa_i}{\lambda_i} \). For comparison with a second multi-index, say \( \lambda \), we use \( (\kappa < \lambda) \iff (\forall i : \kappa_i < \lambda_i) \), where \( < \) can be replaced by \( >, \leq, \) or \( \geq \). For negated comparisons, however, we use \( (\kappa \not< \lambda) \iff (\exists i : \kappa_i \not< \lambda_i) \). The summation symbol \( \sum_{\kappa \geq \lambda} \) stands for summation over all \( \kappa \in \mathbb{N}_0^d \) with \( |\kappa| \leq n \), and can also be interpreted as \( \sum_{\kappa \geq \lambda} \sum_{|\kappa| \leq n} \). We use greek symbols for multi-indices — which may also be used for powers, degrees, etc. — and latin symbols for scalar indices. It should be noted, however, that in case \( d \equiv 1 \) a multi-index degenerates to a scalar. Multi-indices are used to present multivariate algebra with a quasi-univariate notation. We have taken special care to allow for a simple univariate interpretation of the formulas — perhaps on first reading — by interpreting the multi-indices as scalars.

Polynomials and partial derivatives

A \( d \)-variate algebraic monomial in terms of \( x \in \mathbb{R}^d \) is expressed as: \( x^\kappa := \prod_{i=1}^d x_i^{\kappa_i} \), with \( \kappa \) a multi-index. Each \( \kappa_i \) is the degree for dimension \( i \) and \( |\kappa| \) the total degree. The linear space \( P_m^\kappa \subset V \) of algebraic polynomials with maximum total degree \( m \) is defined in terms of \( x \) as \( P_m^\kappa(x) := \text{span}\{x^\kappa | \kappa \in \mathbb{N}_0^d, |\kappa| \leq m\} \). A partial derivative operator is defined analogous to a monomial: \( D^\kappa := \prod_{i=1}^d (\partial/\partial x_i)^{\kappa_i} \), with \( \kappa_i \) the order of differentiation in dimension \( i \) and \( |\kappa| \) the total order. We will use the shorthand \( f^{(\kappa)} := D^\kappa f \).
Taylor expansions

The n-th order truncated Taylor expansion of a function \( f \in V \) around \( x \in X \) may be expressed using operator \( T^n_x : V \to P^n \) as:

\[
T^n_x[f] := \sum_{k=0}^{n} \frac{1}{k!} (x - x)^k f^{(k)}(x).
\]

(1)

Formulation (1) corresponds to the most common perception of a Taylor expansion: as a prediction from a fixed point to its neighborhood. Point \( x \) appears as a subscript to indicate its role as a fixed parameter. In the following it will prove useful to look at the Taylor expansion from the opposite perspective: as a prediction to a fixed point from its neighborhood. This dual perspective calls for a different formulation and notation that we introduce here.

**Definition 2.1.** The n-th order truncated dual Taylor expansion is defined using operator \( \tilde{T}^n_x : V \to V \) as:

\[
\tilde{T}^n_x[f] := \sum_{k=0}^{n} \frac{1}{k!} (x \cdot \cdot \cdot )^k f^{(k)}(\cdot),
\]

(2)

where \( x \) again serves as a fixed parameter.

While the "primal" Taylor expansion \( T^n_x[f] \) produces an approximant to \( f \) on its entire domain \( X \), the dual Taylor expansion \( \tilde{T}^n_x[f] \) assigns to each element of \( X \) an approximation of the value of \( f \) at the specific location \( x \). This property makes the dual Taylor expansion attractive for use in approximation schemes. The duality of \( T^n \) and \( \tilde{T}^n \) is illustrated by the identity \( T^n_x[f](y) = \tilde{T}^n_x[f](x) \). Also note that \( \tilde{T}^n_x[f](x) = f(x) \).

Approximation schemes

Approximation schemes come in all shapes and sizes. In the scope of this paper it is not necessary to discuss details of particular approximation schemes. In fact, for our purpose only one property of a scheme is crucial: its polynomial reproduction order, sometimes referred to as its order of consistency.

We summarize an entire approximation scheme — including, for example, sampling, construction of an approximant, and evaluation at \( x \) — by a single functional that maps the function \( f \) to its approximated value at \( x \).

**Definition 2.2.** An approximator of order \( m \) at \( x \) is a functional \( A^n_m : V \to \mathbb{R} \) that satisfies

\[
A^n_m[p] = p(x) \quad \forall \quad p \in P^m,
\]

(3)

i.e., it reproduces all polynomials of maximum total degree \( m \) at \( x \).

3. Motivation and a simple example

Suppose that we have: (a) an approximation scheme that is represented by an approximator \( A^n_m \) as defined in Definition 2.2, and (b) a corresponding dataset (or sampling strategy) for a function \( f \in V \), such that we can evaluate an approximation \( \hat{f}(x) \) to \( f \) at \( x \) using \( \hat{f}(x) = A^n_m[f] \).

Now, consider the case that the dataset provides \( n \) orders of derivative data that are supplementary, in the sense that they are not used in the approximation scheme. It is then attractive to incorporate this information to improve the approximation. We invoke the truncated dual Taylor expansion (2), and replace \( f \) by \( \tilde{T}^n_x[f] \) prior to the application of \( A^n_m \). This results in an alternative approximation \( \tilde{f}(x) = A^n_m[\tilde{T}^n_x[f]] \) that incorporates all available data.

To assess the resulting procedure we study a simple example based on linear interpolation in a univariate setting (see also [14, Ch. 4], with more details). Let \( z \in [z_0, z_1] \subset \mathbb{R} \) and \( g : [z_0, z_1] \to \mathbb{R} \). The linear interpolation scheme is a first order approximator at \( z \), denoted by \( I_1^L[g] = g(z_0)\phi_0(z) + g(z_1)\phi_1(z) \), with \( \phi_0(z) = (z_1 - z)/(z_1 - z_0) \) and \( \phi_1(z) = (z - z_0)/(z_1 - z_0) \). Application of the procedure suggested above yields the following interpolation scheme

\[
I_1^L[\tilde{T}^1_x[g]] = [g(z_0) + (z - z_0)g'(z_0)]\phi_0(z) + [g(z_1) + (z - z_1)g'(z_1)]\phi_1(z).
\]

(4)

Note that the approximation (4) is now quadratic rather than linear in \( z \). Therefore, it is interesting to see what it does to a second degree polynomial, say \( p(z) = c_0 + c_1 z + c_2 z^2 \), with \( c_0, c_1, c_2 \in \mathbb{R} \) arbitrary constants. It turns out that
\[ I_1^1[p(z)] - p(z) = -c_2(z_1 - z)(z - z_0), \]

while \( I_1^1[p(z)] - p(z) = c_2(z_1 - z)(z - z_0). \) Apparently the suggested procedure does not enhance the accuracy; in fact, the errors are equal in size but opposite in sign. This very observation immediately suggests an attractive alternative: to combine both interpolations. For arbitrary \( g \) this results in a new interpolation scheme

\[ \frac{1}{2}(I_1^1[g] + I_1^1[\tilde{T}_1^1[g]]) = [g(z_0) + \frac{1}{2}(z - z_0)g'(z_0)]\phi_0(z) + [g(z_1) + \frac{1}{2}(z - z_1)g'(z_1)]\phi_1(z), \]

that is able to reproduce \( p \) perfectly.

On the right hand sides of (5) and (4) we see that the dual Taylor expansion has been replaced with an alternative Taylor-like expansion that involves an extra coefficient. We also see that the combination of a first order approximation and a first order expansion results in an approximation of second order accuracy.

In the following sections we will show that this principle can be generalized and that it is possible to combine an \( m \)-th order approximation with an \( n \)-th order Taylor-like expansion to obtain an effective \((m + n)\)-th order accuracy.

4. The reduced dual Taylor expansion

**Definition 4.1.** The \( n \)-th order reduced dual Taylor expansion of the \( m \)-th kind, \( \overline{D}_n^m : V \to V \) is defined as:

\[ \overline{D}_n^m[f] := \sum_{|k| \leq n} \frac{1}{k!} C_{\nu[k]}^m (x - .)^{f^{(\nu[k])}}(.). \]  

with

\[ C_{\nu[k]}^m := \binom{m + n}{m}^{-1} \binom{m + n - |k|}{m}. \]  

The only difference between (6) and (2) is the insertion of the coefficients \( C_{\nu[k]}^m \). Note that \( C_{\nu[k]}^m \leq 1 \) such that all terms in expansion (6) are reduced with respect to those in (2), explaining the expansion’s adjective. It is also worth noting that for \( m = 0 \) the original dual Taylor expansion (2) is obtained. The coefficients \( C_{\nu[k]}^m \) defined in (7) have been chosen specifically for Theorem 5.1 below to hold. To prepare for the proof of Theorem 5.1 we first need to establish the following properties of \( \overline{D}_n^m \):

**Lemma 4.1.** The reduced dual Taylor expansion of a multivariate monomial \((x)^t\) can be expressed as

\[ \overline{D}_n^m[(x)^t] = \sum_{|\lambda| \leq m} B_{\lambda t}^m x^t (x)^{\lambda - t}, \]  

with coefficients \( B_{\lambda t}^m \) satisfying the following properties:

\[ B_{\lambda t}^m = 0 \quad \forall \quad \{t, \lambda | t \leq \lambda \} \]  

\[ B_{\lambda t}^m = 0 \quad \forall \quad \{t, \lambda | (|\lambda| - |t| > m) \wedge (|\lambda| \leq m + n) \} \]  

\[ \sum_{|\lambda| \leq m} B_{\lambda t}^m = 1. \]

**Proof.** Starting from Definition 4.1, expanding the powers of \((x - .)\), and reordering terms gives

\[ \overline{D}_n^m[(x)^t] = \sum_{|k| \leq n} \sum_{|\lambda| \leq m} (-1)^{k - t} C_{\nu[k]}^m \binom{\lambda}{k} \binom{k}{t} x^t (x)^{\lambda - t}. \]

From this we identify the coefficients \( B_{\lambda t}^m \) in (8) as

\[ B_{\lambda t}^m = \sum_{|k| \leq n} (-1)^{k - t} C_{\nu[k]}^m \binom{\lambda}{k} \binom{k}{t}. \]
Next, we invoke a sequence of three established binomial identities (A.1-A.3) for all dimensions successively to arrive at a simpler expression for $B_{x,l}^{mn}$. The result is:

$$B_{x,l}^{mn} = \binom{m+n}{m}^{-1} \binom{m+n-|l|}{n-|l|}$$

(14a)

$$= \binom{m+n}{m}^{-1} \binom{m+n-|l|}{m-|l|+|l|}$$

(14b)

Properties (9) and (10) follow directly from the second and the third binomial coefficient in (14b) respectively. Refer to Section 2 for the definition of negated comparison operators for multi-indices. Property (11) follows from Vandermonde’s identity (A.4) applied to (14a) successively in all dimensions.

5. Raising the approximation order

**Theorem 5.1.** The combination of $n$-th order reduced dual Taylor expansion of the $m$-th kind $\overline{D}_x^{mn}$ (6) and $m$-th order approximator $A_x^m$ (3) yields an effective approximation order of $m+n$:

$$A_x^m[\overline{D}_x^{mn}[p]] = p(x) \quad \forall \ p \in P^{m+n}.\quad (15)$$

**Proof.** From Lemma 4.1, Equations (8) and (10), we see that $\overline{D}_x^{mn}$ maps any monomial $(\cdot)^l$ with $|l| \leq m+n$ to a polynomial of maximum total degree $m$, because the coefficients for the powers $(\cdot)^l$ with $|l| = n-|l| > m$ vanish. Since $\overline{D}_x^{mn}$ is a linear operator, this also holds true for any polynomial of total degree $|l| \leq m+n$. Approximator $A_x^m$ reproduces all polynomials of maximum total degree $m$ (by Definition 2.2). Therefore, the only thing left to prove (15) is to verify that $\overline{D}_x^{mn}[p]$ replicates $p$ at $x$. This is easily checked for an arbitrary monomial, using Equations (8) and (11):

$$\overline{D}_x^{mn}(\cdot)^l(x) = \sum_{|l| \leq m} B_{x,l}^{mn} x^{l-i} = x^l,$$

(16)

which finalizes the proof of Theorem 5.1.

An expression for the approximation error is obtained by [15] and [16] for quasi-interpolation in univariate and multivariate settings respectively. We translate their result to our situation and generalize it using the unspecified approximator $A_x^m$:

$$f(x) - A_x^m[\overline{D}_x^{mn}[f]] = A_x^m\left[\int_0^1 (-1)^m t^m(1-t)^n \frac{d^{m+n+1} f(\cdot + t(x-\cdot))}{dt^{m+n+1}} \right].\quad (17)$$

This expression is obtained starting with a Taylor expansion of $f^{(m)}$ around $x$ in (6). Note that setting $n = 0$ in (17) reveals the error expression for the $A_x^m$ itself. Setting $m = 0$ results in the familiar integral remainder of the Taylor expansion. Error bounds may be derived from (17) for specific choices of $A_x^m$. We refer to [15] and [16] for some specific examples in (quasi-) interpolation.

6. Implementation

The practical consequence of Theorem 5.1 is a generic procedure that we can use to raise the order of multivariate approximation schemes. An approximation scheme for the value of $f$ at $x$ is basically an operator that takes as input a number of samples of $f$, and delivers as output an approximation to $f(x)$. The proposed procedure simply involves replacing the input samples of $f$ by the corresponding samples of $\overline{D}_x^{mn}[f]$, which can be constructed using the supplementary derivative data. The approximation order will then be raised from $m$ to $m+n$. The result is best appreciated by comparing (15) to (3). Note that the supplementary derivative data is assumed to be available at the same locations as the original samples.
Next, we show the effect of the procedure on 2-D scattered data interpolation. We consider a domain \( X = [-0.2, 1.2]^2 \subset \mathbb{R}^2 \), variables \( \mathbf{x} = (x_1, x_2) \in X \) and a function \( f(x) = \cos(2\pi x_1) + x_1 \sin(2\pi x_2) + x_2^2 - x_1 \) to be approximated on \( [0, 1]^2 \subset X \). Data is scattered over \( N \) random node locations \( \{ \mathbf{x}_k \in X \}_{k=1}^N \), and we use shorthands \( f_k := f(x_k) \). We investigate two different approximation techniques for scattered data. To assess their approximation orders we investigate the decrease of the approximation error as a function of the number of nodes \( N \). Since the nodes are spread over 2 dimensions, the average distance between neighbouring nodes scales as \( N^{-1/2} \). Regardless of the approximation technique used we therefore predict the error of the enhanced approximation to be \( O(N^{-(m+n+1)/2}) \), with \( m \) and \( n \) as defined above.

The first approximation \( \tilde{f} \) is constructed using a Delaunay triangulation of the nodes and linear (barycentric) interpolation within each triangle [1]. Let the barycentric weights for location \( \mathbf{x} \) be denoted by \( w_k(\mathbf{x}) \), then the approximation \( \tilde{f} \) can be expressed as

\[
\tilde{f}(\mathbf{x}) = \sum_{k=1}^{N} w_k(\mathbf{x}) f_k.
\] (18)

Note that all \( w_k(\mathbf{x}) \) vanish except for those that correspond to the vertices of the triangle that contains \( \mathbf{x} \). The incorporation of the derivative data now comes down to the replacement

\[ f_k \rightarrow \tilde{f}(\mathbf{x}_k) \] (19)
in (18). Results are shown in Figure 1. Note that \( m = 1 \) in this case.

The second approximation \( \bar{f} \) is constructed using the moving least-squares (MLS) approximation [6]:

\[
\bar{f}(\mathbf{x}) = \left[ \arg\min_{p \in P_m} \sum_{k=1}^{N} \phi(||\mathbf{x} - \mathbf{x}_k||)(p(\mathbf{x}_k) - f_k)^2 \right](\mathbf{x}),
\] (20)

where \( \phi : \mathbb{R}_+^n \rightarrow \mathbb{R}_+^n \) is a positive weighting function that normally peaks at zero and decreases away from it. For our example we use \( \phi(r) = \exp(-Nr^2) \). The incorporation of the derivative data again comes down to the replacement (19) in (20). Results are shown in Figure 2.

7. Remarks

For multivariate approximation we use the concepts of (maximum) total degree for polynomials and total order for approximation schemes and derivatives. These are perhaps the most natural — or at least the most symmetric — generalizations of the univariate concepts of order and degree. As a consequence, both the original and the enhanced approximation schemes have the same approximation order in all directions. Also, the order of derivative information must be the same in all directions. We call this the total-degree approach. An alternative is to apply univariate techniques to all dimensions in succession. This tensor-product approach is well known from, e.g., bilinear and bicubic interpolation and tensor-product B-splines [1, 2, 3] and is very attractive for rectangular data-configurations (grids). It allows the use of different approximation orders in different dimensions. Our procedure can also be used in a purely tensor-product fashion, by simply using the univariate version of (15) in each dimension. Now, like the order \( m \), also the number of (partial) derivatives \( n \) may be varied in each dimension. It should be noted, however, that all (tensor-product) mixed partial derivatives are required as well, and these may not be available. It may then be more attractive to determine the total approximation order for the original tensor-product scheme and use the corresponding multivariate kind of (15).

In our motivation we have referred to \( n \) orders of derivative data that are supplementary, in the sense that they are not used in the approximation scheme. It is not necessary for those supplementary orders to be the lowest orders. It is also possible that some derivative information is accommodated in the original approximation scheme, such as in a Hermite interpolation. In that case, the supplementary derivative information consists of the next \( n \) orders. The total derivative information, however, should be contiguous from order 0 upwards. The procedure does not change at all. The reduced dual Taylor expansion should still be of order \( n \). As the approximation scheme samples the derivatives of (6), the higher order derivative information is introduced automatically.
Figure 1: Scattered data interpolation using barycentric interpolation on Delaunay triangulation. Figures (a)-(c) represent experiment for $N = 75$, i.e. 75 randomly distributed nodes on domain $x \in [-0.2,1.2]^2$. (a) Test function $f(x) = \cos(2\pi x_1) + x_1 \sin(2\pi x_2) + x_2^2 - x_1$, node locations and triangulation. (b) Approximation to $f(x)$ with $m = 1$ and $n = 0$. (c) Approximation to $f(x)$ with $m = 1$ and $n = 1$. (d) Log-log plot of relative rms-error versus $N$. Solid, dashed and dot-dashed curves correspond to $n = 0$, $n = 1$, and $n = 2$ respectively. Convergence rates are as predicted: $O(N^{-(m+n+1)/2})$. 
Figure 2: Scattered data approximation using moving least-squares. Figures (a)-(e) represent experiment for $N = 75$, i.e. 75 randomly distributed nodes on domain $x \in [-0.2, 1.2]^2$. (a) Test function $f(x) = \cos(2\pi x_1) + x_1 \sin(2\pi x_2) + x_2^2 - x_1$ and node locations. (b) MLS weight function $\phi(r) = \exp(-Nr^2)$ (Gaussian). (c) Approximation to $f(x)$ with $m = 1$ and $n = 0$. (d) Approximation to $f(x)$ with $m = 2$ and $n = 0$. (e) Approximation to $f(x)$ with $m = 1$ and $n = 1$. (f) Log-log plot of relative rms-error versus $N$. Black and red curves correspond to $m = 1$ and $m = 2$ respectively; solid, dashed and dot-dashed curves correspond to $n = 0$, $n = 1$ and $n = 2$ respectively. Convergence rates are as predicted: $O(N^{-(m+n+1)/2})$.
Approximation schemes are used for various purposes. In some cases the main goal is accuracy, in other cases smoothness is more important. In that respect it is good to emphasize that our procedure results in a scheme that has (at least) the same order of continuity as the original scheme. Therefore, our procedure is an attractive tool for raising the order of approximation schemes that are designed for smoothness.

8. Conclusion

We have established a generic procedure to raise the effective approximation order of multivariate approximation schemes by incorporating supplementary derivative data. A particular strength of the procedure is that it is easily accommodated in existing schemes, as demonstrated with two examples of scattered data interpolation/approximation. We expect the procedure to be useful immediately in a wide range of applications.

Acknowledgment

A large part of the work presented was carried out while the first author was employed at the Faculty of Civil Engineering and Geosciences, Delft University of Technology, in a project financed by the Netherlands Research Centre for Integrated Solid Earth Science (ISES).

The second author takes part in the research program of the Foundation for Fundamental Research on Matter (FOM), which is financially supported by the Netherlands Organization for Scientific Research (NWO).

Appendix A. Binomial identities

For ease of reference we provide without proof four (standard) binomial identities using symbols that mimic the symbols used in the main text. With $i,k,l,m,n \in \mathbb{N}_0$ we have from [15], p. 174:

\[ \binom{l}{k} \binom{k}{i} = \binom{l}{i} \binom{l-i}{k-i} \]  \hspace{1cm} (A.1)

and from [15], Eqs. 5.25, 5.14 and 5.22 respectively:

\[ \sum_{k=0}^{n} (-1)^{k-i} \binom{m+n-k}{m} \binom{l-i}{k-i} = (-1)^{n-i} \binom{l-i-m-1}{n-i} \]  \hspace{1cm} (A.2)

\[ (-1)^{n-i} \binom{l-i-m-1}{n-i} = \binom{m+n-l}{n-i} \]  \hspace{1cm} (upper negation), (A.3)

\[ \sum_{i=0}^{n} \binom{m+n-l}{n-i} = \binom{m+n}{n} \]  \hspace{1cm} (Vandermonde’s identity). (A.4)

References