On algebraic Stein operators for Gaussian polynomials

EHSAN AZMOODEH\textsuperscript{1} and DARIO GASBARRA\textsuperscript{2} and ROBERT E. GAUNT\textsuperscript{3}

\textsuperscript{1}Department of Mathematical Sciences, Mathematical Sciences Bldg, The University of Liverpool, Liverpool L69 7ZL, Email: ehsan.azmoodeh@liverpool.ac.uk

\textsuperscript{2}Department of Mathematics and Statistics, University of Helsinki, P.O.Box 68 00014, Finland, Email: dario.gasbarra@helsinki.fi

\textsuperscript{3}Department of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, UK, Email: robert.gaunt@manchester.ac.uk

The first essential ingredient to build up Stein’s method for a continuous target distribution is to identify a so-called Stein operator, namely a linear differential operator with polynomial coefficients. In this paper, we introduce the notion of algebraic Stein operators (see Definition 3.2), and provide a novel algebraic method to find all the algebraic Stein operators up to a given order and polynomial degree for a target random variable of the form $Y = h(X)$, where $X = (X_1, \ldots, X_d)$ has i.i.d. standard Gaussian components and $h \in \mathbb{K}[X]$ is a polynomial with coefficients in the ring $\mathbb{K}$. Our approach links the existence of an algebraic Stein operator with null controllability of a certain linear discrete system. A \texttt{MATLAB} code checks the null controllability up to a given finite time $T$ (the order of the differential operator), and provides all null control sequences (polynomial coefficients of the differential operator) up to a given maximum degree $m$. This is the first paper that connects Stein’s method with computational algebra to find Stein operators for highly complex probability distributions, such as $H_{20}(X_1)$, where $H_p$ is the $p$-th Hermite polynomial. Some examples of Stein operators for $H_p(X_1)$, $p = 3, 4, 5, 6$, are gathered in the Appendix and many other examples are given in the Supplementary Information.

Keywords: Stein’s method, Stein operator, Gaussian integration by parts, Malliavin calculus, linear system theory, null controllability, symbolic computation, Hermite polynomials.

1. Introduction

We begin with the following definition that plays a pivotal role in our paper.

\textbf{Definition 1.1.} Let $Y$ be a (continuous) target random variable. We say that a linear differential operator $S = \sum_{\ell=0}^{T} p_{\ell} \partial^{\ell}$ acting on a class $F$ of functions is a polynomial Stein operator for $Y$ if (a) $Sf \in L^1(Y)$, (b) $E[Sf(Y)] = 0$ for all $f \in F$, and (c) the coefficients of $S$ are polynomial. By $\text{PSO}_F(Y)$ we denote the set of all polynomial Stein operators, acting on a class $F$ of functions, for the target random variable $Y$.

In the last decade, polynomial Stein operators have got a lot of attention due to their important role in the Nourdin-Peccati Malliavian–Stein approach Nourdin, I. and Peccati,
This method not only provides a drastic simplification of the classical method of the moments/cumulants, but also allows for quantification of many probabilistic limit theorems that were not possible before. In 1972, Charles Stein introduced a powerful technique for estimating the error in Gaussian approximations. Stein’s method for Gaussian approximation rests on the following fundamental Gaussian integration by parts formula: for $X \sim N(0,1)$ a standard Gaussian random variable,

$$\mathbb{E}[Xf(X) - \partial f(X)] = 0$$  \hfill (1.1)

for all absolutely continuous functions $f : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}|\partial f(Y)| < \infty$. Here $\partial = \frac{d}{dx}$ is the usual differentiation operator. This formula leads to the so-called Stein equation:

$$xf(x) - \partial f(x) = h(x) - \mathbb{E}[h(X)],$$  \hfill (1.2)

where the test function $h$ is real-valued. It is straightforward to verify that $f(x) = -e^{-x^2/2} \int_{-\infty}^{x} \{h(t) - \mathbb{E}[h(X)]\}e^{-t^2/2} dt$ solves (1.2), and bounds on the solution and its derivatives in terms of the test function $h$ and its derivatives are given in Chen, L. H. Y., Goldstein, L. and Shao, Q.-M. (2011); Döbler, C., Gaunt, R. E. and Vollmer, S. J. (2017). Evaluating both sides of (1.2) at a random variable $W$ and taking expectations gives

$$\mathbb{E}[Wf(W) - \partial f(W)] = \mathbb{E}[h(W)] - \mathbb{E}[h(X)].$$  \hfill (1.3)

Thus, the problem of bounding the quantity $|\mathbb{E}[h(W)] - \mathbb{E}[h(X)]|$ has been reduced to bounding the left-hand side of (1.3). A detailed account of Stein’s method for Gaussian approximation and some of its numerous applications throughout the mathematical sciences are given in the monograph Stein, C. (1986) and the books Chen, L. H. Y., Goldstein, L. and Shao, Q.-M. (2011); Nourdin, I. and Peccati, G. (2012).

One of the advantages of Stein’s method is that the above procedure can be extended to treat many other distributional approximations. In adapting the method to a new continuous distribution, the first step is to find a suitable analogue of the integration by parts formula (1.1). For a target random variable $Y$, this amounts to seeking a Stein operator $S$ acting on a class of functions $F$ such that $Sf \in L^1(Y)$ and $\mathbb{E}[Sf(Y)] = 0$.

For continuous distributions, $S$ is typically a differential operator, although some operators in the literature are integral or fractional Arras, B. and Houdré, C. (2019); Xu, L. (2019). As noted by Barbour, A. D. (1990), for a given distribution, there are infinitely many Stein operators. A common approach to sifting through the available options is to restrict to $T$-th order polynomial Stein operators (see Definition 1.1) of the form

$$Sf(y) = \sum_{t=0}^{T} p_t(y) \partial^t f(y)$$  \hfill (1.4)

in which the coefficients belongs to the polynomial ring $\mathbb{K}[y]$ in the single variable $y$. (In this paper, $\partial^0 \equiv I$, the identity operator. For ease of exposition, we abuse notation and write $y$ in place of $yI$.) In addition to their utility in the Malliavin-Stein method, such Stein operators are amenable to the various coupling techniques used in the implementation of...
Stein’s method to derive distributional approximations, and as such the vast majority of
differential Stein operators used in the literature take this form; for an overview see
Gaunt, R. E., Mijoule, G. and Swan, Y. (2019); Ley, C., Reinert, G. and Swan, Y. (2017).
The density method Stein, C. (1986); Ley, C. and Swan, Y. (2013); Ley, C., Reinert, G. and Swan, Y. (2017); Stein, C. et al. (2004) leads to first order Stein operators, which
are polynomial if the log derivative of the density is a rational function; this approach
provides first order polynomial Stein operators for, amongst others, target distributions
which belong to the Pearson family Schoutens, W. (2001) or which satisfy a diffusive
assumption Döbler, C. (2015); Kusuoka, S. and Tudor, C. A. (2012). Another method that
naturally leads to first order Stein operators is the generator method of Barbour, A. D.
(1990); Götz, F. (1991). However, it is often necessary to consider higher order operators;
second order polynomial operators are needed for the Laplace Pike, J. (2014), PRR Peköz,
E., Röllin, A. and Ross, N. (2013) and variance-gamma Gaunt, R. E. (2014) distributions,
for example. This is because the densities of these distributions satisfy second order
differential equations with polynomial coefficients. In recent years, techniques have been
developed for obtaining Stein operators in increasingly complex settings, such as the
iterated conditioning argument for deriving Stein operators for products of a quite general
class of distributions Gaunt, R. E. (2017, 2018); Gaunt, R. E., Mijoule, G. and Swan, Y.
(2019, 2020) and the Fourier/Malliavin calculus approach used to obtain Stein operators
for linear combinations of gamma random variables Arras, B. et al. (2019, 2020).

However, there remain many important distributions for which Stein’s method has
not been adapted to. One identified by Peccati, G. (2014) to be of particular importance
are those of the form $P(X)$, where $X \sim N(0,1)$ and $P$ is a polynomial of degree strictly
greater than 2. The case $P$ is the $p$-th Hermite polynomial, $H_p(x) = (-1)^p e^{x^2/2} \partial_x^p(e^{-x^2/2})$,
is of particular interest, due to their fundamental role in Gaussian analysis and Malliavin
calculus. More importantly, they appear as target distributions in the asymptotic theory
of $U$-statistics, see (Koroljuk, V. S., Borovskich and Yu. V., 1994, Section 4.4) and (Lee,
A. J., 1990, Chapter 3). For example, taking the kernel $\psi(x_1, \ldots, x_p) = x_1x_2 \ldots x_p$, the
limit of the corresponding $U$-statistics is $H_p(X)$, see p. 87 in the latter reference. Since
the seminal paper Nourdin, I. and Peccati, G. (2009), the Malliavin-Stein method has
been used to derive quantitative limits theorems for a wide class of laws (see, for example,
Eden, R. and Viens, F. (2013); Eden, R. and Viquez, J. (2015)), and has been particularly
successful for target distributions from the first or second Wiener chaos (e.g., Arras, B.
et al. (2019, 2017); Azmoodeh, E. and Gasbarra, D. (2018); Azmoodeh, E., Peccati, G.
and Poly, G. (2015); Döbler, C. and Peccati, G. (2018); Eichelsbacher, P. and Thäle, C.
(2014); Nourdin, I., Nualart, D. and Peccati, G. (2016); Nourdin, I. and Peccati, G. (2009);
Nourdin, I., Peccati, G. and Réveillac, A. (2010); Nourdin, I. and Peccati, G. (2015);
Nourdin, I. and Poly, G. (2012)), but little is known about convergence to targets from
higher order chaoses. In the Stein’s method literature, the only contribution is the modest
one of Gaunt, R. E. (2019) in which elementary but involved manipulations led to fifth
and third order polynomial Stein operators for $H_5(X)$ and $H_3(X)$, respectively, representing
a first step towards the goal of using the Malliavin-Stein method to understand convergence
towards elements of higher order Wiener chaoses. In this paper we make progress that far
In this work, we consider the multivariate Gaussian polynomial $Y = h(X)$, where $X = (X_1, \ldots, X_d)$ has independent and identically distributed (i.i.d.) standard Gaussian components and $h \in \mathbb{K}[X] = \mathbb{K}[X_1, \ldots, X_d]$, $\mathbb{K} = \mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$. For any target random variable $Y = h(X)$ of the above form, we show, in the formalism of linear system theory, that in Gaussian space, finding Stein operators of the type (1.4) with polynomial coefficients $p_t(y) \in \mathbb{K}[Y]$ is equivalent to solving a null controllability problem.

Controllability is a fundamental concept in mathematical control theory. Consider a (finite-dimensional) discrete-time linear time-invariant system $x_{k+1} = Ax_k + Bu_k$, $k \in \mathbb{N}$ where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the input, and $A$ and $B$ are matrices of suitable size. The controllability property of such a system refers to the fact that any initial state can be steered to any final state by choosing the input appropriately. This notion was introduced by Kalman, R. E. (1960) and was extensively studied by Kalman himself and many others; see Kalman, R. E. (1963); Kailath, T. (1980); Sontag, E. D. (1998) and the survey Klamka, J. (1993). Controllability has many applications in control theory and systems theory, as well as in industrial and chemical process control, reactor control, control of electric bulk power systems, aerospace engineering and quantum systems theory.

The significance of formulating the problem of finding polynomial Stein operators for Gaussian polynomials as a null controllability lies in the fact that the problem of determining the controllability of a discrete-time linear system is well understood in the literature. The classical Kalman’s rank test $\text{rank} [B \ AB \ \cdots \ A^{n-1}B] = n$ is a necessary and sufficient condition for the controllability of a linear finite dimensional discrete system. An alternative characterisation, which is sometimes called the Popov-Belevitch-Hautus (PBH) test, presents an equivalent condition in terms of the eigenmodes of the system Antsaklis, P. J. and Michel, A. N. (2007); Bourlès, H. (2010). In Remark 3.5, we present the algorithm used in this paper for finding a null control sequence of a discrete-time linear system that can be implemented efficiently in modern computer algebra packages.

In Definition 3.2, we introduce the notion of algebraic Stein operators, a subclass of the class of polynomial Stein operators, which coincides with the class of polynomial Stein operators in the case $d = 1$ (see Proposition 3.2). For a given target $Y = h(X)$, the above procedure can be implemented as an algorithm in modern computer algebra packages that can find all algebraic Stein operators for $Y = h(X)$ up to a given order $T$ and maximal degree of polynomial coefficients $m = \max_{0 \leq t \leq T} \deg(p_t(y))$. We provide an efficient MATLAB code at https://github.com/gasbarra/Algebraic-Stein-Equations.

In dimension $d = 1$, we use the method to study the Stein operators for targets $Y = H_p(X)$. The ‘simplest’ Stein operators for $p = 1, \ldots, 6$ are displayed in Appendix A. The Supplementary Information contains many more examples. For $p \geq 7$, the numerical coefficients in the Stein operators become too large to present in this paper, but we give some useful summaries of these Stein operators in Table 1. It should be clear that, apart from the first few cases, it would be very hard to derive these Stein operators solely by bare-hands heroic calculations. This paper therefore constitutes not only the first connection between Stein’s method and control theory, but also the first study that deeply connects Stein’s method with computer algebra programming to find Stein operators for highly complex probability distributions, such as $H_{20}(X)$.
In addition to being able to find Stein operators that are out of reach of existing methods, our connection between Stein’s method and null controllability has a feature not seen before in the Stein’s method literature. As our algorithm finds all algebraic Stein operators for a given target distribution \( Y = h(X) \) up to a given order \( T \) and polynomial degree \( m \), we have confidence (at least in the case \( d = 1 \)) as to what polynomial Stein operators for that distribution are ‘simplest’ in the sense of having minimal order \( T \) or polynomial degree \( m \). For example, by running our MATLAB code to obtain polynomial Stein operators for \( H_3(X) \) and \( H_4(X) \) we know (with the aid of Proposition 4.1) that the fifth and third order operators that Gaunt, R. E. (2019) obtained (see (A.1) and (A.2)) are ‘simplest’ in the sense of having minimum possible maximal degree 2 and have the minimum possible order \( T \) for such operators, which was not obvious from the analysis of Gaunt, R. E. (2019). The complexity of Stein operators for \( H_p(X) \) only increases as \( p \) increases (see Table 1), with the implication being that there is a significant increase in difficulty in applying the Malliavin-Stein method for target distributions \( H_p(X) \) for \( p \geq 3 \). Another insight we gain from the outputs of our algorithm is that for \( p \geq 3 \) there is not a unique ‘simplest’ Stein operator in that the Stein operator with smallest possible order \( T \) (degree \( m \)) does not have smallest possible maximal polynomial degree \( m \) (order \( T \)). It remains to be seen whether the Stein operators with minimal order \( T \) or minimal degree \( m \) are more tractable when used in the Malliavin-Stein method. Finally, we stress that the method of converting the problem of finding Stein operators into a null controllability problem applies more generally than even the Gaussian polynomial setting detailed in this paper. We expect this to be explored in future research.

The paper is organised as follows. In Section 2, we review some fundamental Malliavin calculus operators and formulas on \( \mathbb{K} [x_1, \ldots, x_d] \) needed in the paper. Section 3 is the heart of the paper. In Section 3.1, we introduce the notion of a Stein chain, and make a connection between the existence of a Stein chain and polynomial Stein operators for Gaussian polynomials. We present a method for validating Stein chains in Section 3.2. In Section 3.3, we formulate the problem of finding Stein operators for Gaussian polynomials as a null controllability problem. A description of the implementation of the null controllability approach in MATLAB is given in the Supplementary Information (SI).

In Section 4, we focus on applications to Gaussian Hermite polynomials. We provide a description of the highest order coefficient of polynomial Stein operators for \( H_p(X) \) in Section 4.1. In Section 4.2, we summarise some interesting features of the Stein operators for \( H_p(X) \) that are obtained by our MATLAB code; examples of the Stein operators are given in Appendix A and the SI. Finally, an auxiliary lemma is stated and proved in Appendix B.

Note on the class of functions \( \mathcal{F} \): Consider the polynomial Stein operator \( S = \sum_{t=0}^{T} p_t(y) \partial^t \), with \( \max_{0 \leq t \leq T} \deg(p_t(y)) = m \), for the target random variable \( Y \), supported on \( I \subseteq \mathbb{R} \). In this paper, the class of functions on which our Stein operators act is \( \mathcal{F}_{S,Y} \), defined as the set of functions \( f \in C^T(I) \) such that \( \mathbb{E}[Y^j f^{(t)}(Y)] < \infty \) for all \( t = 0, \ldots, T \) and \( j = 0, \ldots, m \). We write \( PSO(Y) \) as shorthand for \( PSO_{\mathcal{F}_{S,Y}}(Y) \). We do not claim that this is the largest class of functions on which the Stein operators of this paper act, but the class is large enough for practical purposes and guarantees that in all of our proofs the crucial property \( \mathbb{E}[S f(Y)] = 0, \forall f \in \mathcal{F} \), holds.
2. Malliavin operators on \( \mathbb{K}[x_1, \ldots, x_d] \)

For the scope of our paper, it is enough to define Malliavin calculus operators on \( \mathbb{K}[x_1, \ldots, x_d] \) (see below for definitions) algebraically, without discussing their functional analytic extensions. For a state-of-the-art exposition of Malliavin calculus in full generality see Nourdin, I. and Peccati, G. (2012); Nualart, D. and Nualart, E. (2018).

2.1. Algebraic preliminaries

We denote by \( \mathbb{K}[x_1, \ldots, x_d] \) the commutative ring of polynomials in the variables \( x_1, \ldots, x_d \) with coefficients in \( \mathbb{K} \), equipped with the usual addition and multiplication.

**Definition 2.1 (Ideal).** A subset \( J \subseteq \mathbb{K}[x_1, \ldots, x_d] \) is an ideal when

1. \( g_1, g_2 \in J \Rightarrow (g_1 + g_2) \in J; \)
2. \( f \in \mathbb{K}[x_1, \ldots, x_d], \ g \in J \Rightarrow fg \in J. \)

**Definition 2.2.** The ideal generated by \( A \subseteq \mathbb{K}[x_1, \ldots, x_d] \) is the smallest ideal of \( \mathbb{K}[x_1, \ldots, x_d] \) containing \( A \), denoted by

\[
\langle A \rangle = \bigcap_{\text{ideal } J \supseteq A} J = \left\{ \sum_{\ell=1}^{n} f_\ell g_\ell : f_\ell \in \mathbb{K}[x_1, \ldots, x_d], g_\ell \in A, n \in \mathbb{N} \right\}.
\]

When \( J = \langle g_1, \ldots, g_n \rangle \) we say that the ideal is finitely generated.

**Remark 2.1.** The ring \( \mathbb{K}[x_1] \) is a principal ideal domain, meaning that every ideal \( I \subseteq \mathbb{K}[x_1] \) is principal (can be generated by one element). This useful fact gives a precise description of the structure of the highest order polynomial coefficients in Stein operators, see Proposition 4.1.

2.2. One-dimensional case

The major application of our algebraic method is to present Stein operators for univariate Gaussian Hermite polynomials. We start with a purely algebraic presentation of Malliavin operators in dimension \( d = 1 \).

**Definition 2.3.** In the univariate case, the Malliavin derivative \( D \), the divergence \( \delta \) and its pseudo-inverse \( \delta^{-1} \) are defined as linear mappings acting on the polynomial ring \( \mathbb{K}[x] = \mathbb{K}[x_1] \), with

\[
D x^n = \partial x^n = nx^{n-1}, \quad \delta x^n = (x - \partial) x^n = x^{n+1} - nx^{n-1}, \quad \text{and}
\]

\[
\delta^{-1} 1 = 0, \quad \delta^{-1} x = 1, \quad \delta^{-1} x^n = x^{n-1} + (n - 1) \delta^{-1} x^{n-2} = \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \frac{(n-1)!!}{(n-1-2k)!!} x^{n-1-2k},
\]

where \( n!! \) denotes the double factorial.
Proposition 2.1. Let \( n \in \mathbb{N}_0 \). It holds that \( \delta^{-1} \delta x^n = x^n \). Moreover, \( \delta \delta^{-1} x^n = x^n - \mathbb{E}[X^n] \), where \( X \sim N(0,1) \) has moment sequence \( \mathbb{E}[X^n] = (n-1)! \) when \( n \in 2\mathbb{N} \) and \( \mathbb{E}[X^n] = 0 \) otherwise.

Lemma 2.1 (Gaussian integration by parts). Let \( X \sim N(0,1) \). For \( f, g \in \mathbb{K}[x], \)

\[
\mathbb{E}[f(X)Dg(X)] = \mathbb{E}[f(X)\partial g(X)] = \mathbb{E}[g(X)\delta f(X)].
\]

Proof. By linearity, it suffices to consider the monomials \( f(x) = x^n, g(x) = x^m \). We have

\[
\mathbb{E}[X^n DX^m] = (m-1)\mathbb{E}[X^{n+m-1}] = \begin{cases} (n + m - 2)!!(m - 1), & (n + m) \text{ odd}, \\ 0, & (n + m) \text{ even}, \end{cases}
\]

which coincides with \( \mathbb{E}[X^m \delta X^n] = \mathbb{E}[X^{m+n+1}] - n\mathbb{E}[X^{m+n-1}] \).

Definition 2.4. We define the univariate Hermite polynomials as \( H_0(x) = 1, H_1(x) = x, \) and

\[
H_n(x) = \delta H_{n-1}(x) = (\delta^n 1)(x), \quad n \geq 2.
\]

Proposition 2.2 (Properties of Hermite polynomials). \( \delta \delta^{-1} H_n(x) = H_{n-1}(x)1(n > 0) \).

1. \( H_0(x), \ldots, H_n(x) \) are monic polynomials spanning \( \mathbb{K}_n[x] \) (the ring of polynomials of maximum degree \( n \)).

2. \( \delta^{-1} H_n(x) = H_{n-1}(x)1(n > 0) \).

3. Hermite polynomials are orthogonal in the sense that for \( X \sim N(0,1) \) it holds that

\[
\mathbb{E}[H_m(X)H_n(X)] = \mathbb{E}[H_m(X)(\delta^n 1)(X)] = \mathbb{E}[\partial_n H_m(X)] = \begin{cases} n!, & n = m, \\ 0, & \text{otherwise}. \end{cases}
\]

4. For \( n \geq 1 \), \( DH_n(x) = \partial H_n(x) = nH_{n-1}(x) \).

5. For every \( 0 \leq m \leq n \), and \( X \sim N(0,1) \),

\[
\mathbb{E}[X^n H_m(X)] = 1((n - m) \in 2\mathbb{N}) \frac{n!(n - m - 1)!!}{(n - m)!}.
\]

Definition 2.5. For a given (target) polynomial \( y = h(x) \), with \( h \in \mathbb{K}[x] \), we introduce the linear operator \( \Gamma_y \) acting on \( \mathbb{K}[x] \) as

\[
\Gamma_y(f(x)) = Dh(x)\delta^{-1}(f(x)) = \partial h(x)\delta^{-1}(f(x)) \in \langle \partial h(x) \rangle.
\]

Lemma 2.2. Let \( X \sim N(0,1) \). Consider \( Y = h(X) \) where \( h \in \mathbb{K}[x] \). Assume that \( f, g \in \mathbb{K}[x] \). Then

\[
\mathbb{E}[g(Y)f(X)] - \mathbb{E}[g(Y)]\mathbb{E}[f(X)] = \mathbb{E}[\partial g(Y)\Gamma_y(f(X))].
\]

Proof. Using Gaussian integration by parts formula Lemma 2.1, we can write

\[
\mathbb{E}[g(Y)f(X)] - \mathbb{E}[g(Y)]\mathbb{E}[f(X)] = \mathbb{E}[g(h(X))\delta \delta^{-1}(f(X))] = \mathbb{E}[\partial (g \circ h)(X) \delta^{-1} f(X)]
\]

\( \square \)
2.3. Multidimensional case

**Definition 2.6.** In the multivariate setting, we define the Malliavin operators as follows.

(a) The Malliavin derivative maps a polynomial $g \in \mathbb{K}[x_1, \ldots, x_d]$ into its gradient:

$$D\bullet g = (D_k g)_{1 \leq k \leq d} = \nabla g = (\partial_{x_1} g(x_1, \ldots, x_d), \ldots, \partial_{x_d} g(x_1, \ldots, x_d)) \in \mathbb{K}[x_1, \ldots, x_d]^d.$$ 

(b) The divergence $\delta$ maps in the opposite direction a polynomial vector $f(x) = f_\bullet(x_1, \ldots, x_d) = (f_1(x_1, \ldots, x_d), \ldots, f_d(x_1, \ldots, x_d)) \in \mathbb{K}[x_1, \ldots, x_d]^d$ into the polynomial

$$\delta f_\bullet(x) = \sum_{k=1}^d \delta_k f_k(x) = \sum_{k=1}^d (f_k(x)x_k - \partial_{x_k} f_k(x)) \in \mathbb{K}[x_1, \ldots, x_d],$$

where $\delta_k$ denotes the univariate divergence operator acting on the $k$-th coordinate.

**Proposition 2.3.** For $X \sim N(0, I_d)$, a standard $d$-dimensional Gaussian random vector with covariance matrix the $d \times d$ identity matrix $I_d$, we have the multivariate Gaussian integration by parts formula

$$\mathbb{E}[g(X)\delta f_\bullet(X)] = \mathbb{E}[(D_\bullet g(X), f_\bullet(X))_{\mathbb{R}^d}].$$

Here, and elsewhere, $(\cdot, \cdot)_{\mathbb{R}^d}$ denotes the usual inner product on Euclidean space $\mathbb{R}^d$.

In order to define a pseudo-inverse of $\delta$ in the multivariate setting, we use the multivariate Hermite polynomials instead of the monomial basis.

**Definition 2.7.**

(a) The Ornstein-Uhlenbeck operator, which maps $\mathbb{K}[x_1, \ldots, x_d]$ into itself, is defined as $L := -\delta D$.

(b) For a $d$-dimensional multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$, the multivariate Hermite polynomials

$$H_\alpha(x) = H_\alpha(x_1, \ldots, x_d) = \prod_{k=1}^d H_{\alpha_k}(x_k)$$

are eigenfunctions of $L$ with respective eigenvalues $-|\alpha| := -\sum_{k=1}^d \alpha_k$. It also follows that $\mathbb{E}[H_\alpha(X)H_\beta(X)] = \mathbf{1}(\alpha = \beta) \prod_{k=1}^d \alpha_k!$.

(c) The linear mapping $L^{-1}$ operates on the multivariate Hermite polynomials as

$$L^{-1}H_\alpha(x) = \begin{cases} 0, & \text{when } |\alpha| = 0, \\ -|\alpha|^{-1}H_\alpha(x), & \text{otherwise}. \end{cases}$$

(d) A pseudo-inverse of $\delta$ is defined as $\delta^{-1} = -DL^{-1}$, with $\delta^{-1}1 = 0$, and, for $|\alpha| > 0$,

$$\delta^{-1}H_\alpha(x) = |\alpha|^{-1}DH_\alpha(x) = \left(\frac{\alpha_1}{|\alpha|}H_{\alpha-e_1}(x), \ldots, \frac{\alpha_d}{|\alpha|}H_{\alpha-e_d}(x)\right),$$

where as usual $(e_i : i = 1, \ldots, d)$ denote the standard basis for the Euclidean space $\mathbb{R}^d$. It follows by definition that $\delta\delta^{-1}H_\alpha(x) = H_\alpha(x)\mathbf{1}(|\alpha| > 0)$. 
On algebraic Stein operators for Gaussian polynomials

**Definition 2.8.** The $n$-th polynomial chaos in the variables $x_1, \ldots, x_d$ is the linear subspace $H_n[x] = H_n[x_1, \ldots, x_d] \subseteq K_n[x] = K_n[x_1, \ldots, x_d]$ generated by the multivariate Hermite polynomials $H_\alpha(x)$ with $|\alpha| = n$, and the decomposition

$$K_n[x] = K + H_1[x] + \cdots + H_n[x]$$

is orthogonal with respect to the $d$-dimensional standard Gaussian measure.

**Definition 2.9.** In the multivariate case, with (target) polynomial $y = h(x_1, \ldots, x_d)$, and $h \in K[x_1, \ldots, x_d]$, we define the linear Gamma operator $\Gamma_y$ as: $(x = (x_1, \ldots, x_d))$

$$\Gamma_y(f(x)) = (Dy, \delta^{-1} f(x))_{\mathbb{R}^d} = (Dh(x), -DL^{-1} f(x))_{\mathbb{R}^d}. \quad (2.2)$$

For example, $\Gamma_y(1) = 0$ and for a multivariate Hermite polynomial

$$\Gamma_y(H_\alpha(x)) = \sum_{k=1}^{d} \frac{\alpha_k}{|\alpha|} H_{\alpha-e_k}(x) \partial_{x_k} h(x).$$

Note that $\Gamma_y$ maps $K[x]$ into the ideal $\langle \nabla h(x) \rangle = \langle \partial_{x_1} h(x), \ldots, \partial_{x_d} h(x) \rangle$ generated by the gradient of the target polynomial.

**Lemma 2.3.** For $X = (X_1, \ldots, X_d) \sim N(0, I_d)$, $Y = h(X)$, with $h \in K[x_1, \ldots, x_d]$, and every $f \in K[x_1, \ldots, x_d], g \in K[Y]$

$$E[g(Y)f(X)] = E[g(Y)]E[f(X)] + E[\partial g(Y) \Gamma_y(f(X))]. \quad (2.3)$$

### 3. An algebraic formalism of Stein operators

Throughout this section, we adopt the following setting. Let $d \geq 1$. Assume $X = (X_1, \ldots, X_d)$, where $X_1, \ldots, X_d$ are i.i.d. standard Gaussian random variables. Choose a polynomial $h \in K[x] = K[x_1, \ldots, x_d]$. We consider a polynomial target random variable $Y$ in the Gaussian variates, namely that,

$$Y = h(X) = h(X_1, \ldots, X_d). \quad (3.1)$$

Without loss of generality, we also assume that $E[Y] = 0$ (this assumption is always possible with a shift). Furthermore, the set $K[y] \cong K[h(x)]$ of all polynomials in the variable $y$ composed with the target polynomial $h(x)$ is a subring of $K[x_1, \ldots, x_d]$. Denote by $H(Y) = L^2(\Omega, \sigma(Y), \gamma_d)$ the Hilbert space of all $Y$-measurable, square integrable random variables with respect to the $d$-dimensional standard Gaussian measure $\gamma_d$. Also, $H(Y)^\perp$ stands for the orthogonal complement of $H(Y)$.
3.1. Forward Stein chain

We begin with the notion of a Stein chain. Our definition is quite abstract, but simply speaking, a Stein chain for the target random variable $Y$ is the sequence of polynomial coefficients of a Stein operator for $Y$ (when it exists); hence its name. This connection is made in Proposition 3.1. A benefit of this abstract definition is that it allows us to define the notion of an algebraic Stein chain in a natural manner. The distinction between general and algebraic Stein chains is important, because Stein operators whose polynomial coefficients form an algebraic Stein chain are easier to find; see Remark 3.2, item (i).

**Definition 3.1 (Forward Stein chain).** Suppose that the target random variable $Y$ has the Gaussian polynomial form given by relation (3.1).

(a) A (general) Stein chain of length $T \geq 1$ for $Y$ is a sequence $(p_t(y) \in \mathbb{K}[y] : t = 0, \ldots, T)$ such that: $g_T + p_T(Y) \in H(Y)^\perp$ where the sequence of Gaussian polynomials $g = (g_t \in \mathbb{K}[X_1, \ldots, X_d] : t = 0, \ldots, T)$ is recursively defined via

\[
\begin{align*}
    g_0 &= 0, \\
    g_t &= \Gamma_Y(g_{t-1} + p_{t-1}(Y)), \quad t = 1, \ldots, T, \\
\end{align*}
\]

where the Malliavin Gamma mapping $\Gamma_Y$ is defined in (2.1) and (2.2) when $d = 1$ and $d \geq 2$, respectively, and moreover $\mathbb{E}[g_t] = -\mathbb{E}[p_t(Y)]$ for $t = 0, \ldots, T$.

Let $SC(Y,T)$ denote the set of all Stein chains of length $T$, and let $SC(Y) = \bigcup_{T \geq 1} SC(Y,T)$ stand for the set of all Stein chains of finite length.

(b) An algebraic Stein chain of length $T \geq 1$ for $Y$ is a Stein chain $(p_t(y) \in \mathbb{K}[y] : t = 0, \ldots, T)$ such that: $g_T + p_T(Y) \in \{0\}$ almost surely. Let $ASC(Y,T)$ denote the set of all algebraic Stein chains of length $T$, and let $ASC(Y) = \bigcup_{T \geq 1} ASC(Y,T)$ stand for the set of all algebraic Stein chains of finite length.

**Remark 3.1.** (i) The moment property $\mathbb{E}[g_t] = -\mathbb{E}[p_t(Y)]$ for $t = 0, \ldots, T$ in Definition 3.1 is immaterial, and can always be assumed due the fact that the linear map $\Gamma_Y$ does not see constants, namely, $\Gamma_Y(f(X) + c) = \Gamma_Y(f(X))$ for any constant $c$.

(ii) We stress that the subring $\mathbb{K}[Y]$ is not stable under the linear mapping $\Gamma_Y$, meaning $\Gamma_Y(\mathbb{K}[Y]) \not\subseteq \mathbb{K}[Y]$. For example, a common candidate – from the Malliavin-Stein perspective – for the first polynomial entry in the Stein chain is $p_0(Y) = cY$, $c \in \mathbb{K}$. It can be readily seen that with the choices $Y = H_p(X)$, $p \geq 3$, and $p_0(Y) = cY$ we have $g_1 \not\in \mathbb{K}[Y]$. Therefore, the most notable feature of a Stein chain or an algebraic Stein chain is that the final output of the linear machine (3.2), either conditionally on the target variable $Y$ or without, is an element in the subring $\mathbb{K}[Y]$.

(iii) Note that by the definition of the $\Gamma_Y$ operator, all the intermediate Gaussian polynomials

\[
g_t = g_t(X_1, \ldots, X_d) \in \langle \nabla h(X_1, \ldots, X_d) \rangle, \quad \forall t = 0, \ldots, T.
\]

More importantly, when the Stein chain is algebraic, $g_T \in \mathbb{K}[Y] \cap \langle \nabla h(X_1, \ldots, X_d) \rangle$. On the other hand, $I = \mathbb{K}[Y] \cap \langle \nabla h(X_1, \ldots, X_d) \rangle$ is an ideal of the subring $\mathbb{K}[Y]$.
and therefore it can be generated by one element (see Remark 2.1). In dimension $d = 1$, for the targets of the form of a Gaussian Hermite polynomial, we present a prototypical element that generates the corresponding ideal, see Proposition 4.1 for details.

Let us continue with the following basic fact that tells us that the existence of a Stein chain of length $T$, in a natural way, leads to a polynomial Stein operator of order $T$ for the target random variable $Y$.

**Proposition 3.1.** Suppose that all the assumptions of Definition 3.1 prevail. Then, the following statements hold:

(a) For every $T \geq 1$, $ASC(Y,T) \subseteq SC(Y,T)$. Moreover, $ASC(Y) \subseteq SC(Y)$.

(b) Let $(p_t(y) \in \mathbb{K}[y] : t = 0, \ldots, T)$ be a forward Stein chain for the target random variable $Y$. Then $S = \sum_{t=0}^{T} p_t \partial_t \in PSO(Y)$. Conversely, if $S = \sum_{t=0}^{T} p_t \partial_t \in PSO(Y)$, then the sequence of polynomials $(p_t(y) \in \mathbb{K}[y] : t = 0, \ldots, T) \in SC(Y,T)$ is a forward Stein chain for $Y$ of length $T$.

**Proof.** (a) This is clear because $0 \in H(Y)^\perp$. (b) Using iteratively the integration by parts formula (2.3), along with the moment property $\mathbb{E}[g_t + p_t(Y)] = 0$, for $t = 0, \ldots, T$, in Definition 3.1 (in particular, $\mathbb{E}[p_0(Y)] = g_0 = 0$), for every $f \in \mathcal{F}$, we can write

$$
\mathbb{E} \left[ (p_0(Y) - \mathbb{E}[p_0(Y)]) f(Y) \right] = \mathbb{E} \left[ (\partial f(Y) \Gamma_Y (p_0(Y) + g_0) \right]$

$$= - \mathbb{E}[p_1(Y) \partial f(Y)] + \mathbb{E} \left[ \partial^2 f(Y) \Gamma_Y (p_1(Y) + g_1) \right] =$$

$$- \mathbb{E}[p_1(Y) \partial f(Y)] - \mathbb{E}[p_2(Y) \partial^2 f(Y)] \cdots - \mathbb{E}[p_{T-1}(Y) \partial^{T-1} f(Y)] + \mathbb{E} \left[ \partial^T f(Y) \Gamma_T (p_{T-1}(Y) + g_{T-1}) \right].$$

Hence, $\mathbb{E} \left[ \sum_{t=0}^{T} p_t(Y) \partial_t f(Y) \right] = \mathbb{E}[p_0(Y)]\mathbb{E}[f(Y)] = 0$, and therefore, the operator $S = \sum_{t=0}^{T} p_t \partial_t$ is a polynomial Stein operator for the target random variable $Y$. For the other direction, let $S = \sum_{t=0}^{T} p_t \partial_t \in PSO(Y)$ be a polynomial Stein operator for $Y$. Then the condition $\mathbb{E}[Sf(Y)] = 0$ for all $f \in \mathcal{F}$ together with an iterative use of the integration by parts formula yields that $\mathbb{E} \left[ \partial^T f(Y) (p_T(Y) + g_T) \right] = 0$ for every $f \in \mathcal{F}$. Hence, using a standard density argument we can conclude that $\mathbb{E} \left[ p_T(Y) + g_T \right] = 0$, and so $g_T + p_T(Y) \in H(Y)^\perp$.

Proposition 3.1, item (b), provides a neat correspondence between $PSO(Y)$ and $SC(Y)$. This useful fact in addition to the novel notion of an algebraic Stein chain introduced in Definition 3.1, item (b), directs us to the notion of an algebraic polynomial Stein operator that is the cornerstone of our paper.
Definition 3.2 (Algebraic polynomial Stein operator). Let \( d \geq 1 \), and suppose that the target random variable \( Y \) has the Gaussian polynomial form given by relation (3.1). Let \( T \geq 1 \), and \( S = \sum_{t=0}^{T} p_t \partial^t \in \text{PSO}(Y) \). We say that \( S \) is an algebraic polynomial Stein operator of length \( T \) for target random variable \( Y \) when the sequence of polynomials \((p_t(y) \in \mathbb{K}[y] : t = 0, \ldots, T) \in \text{ASC}(Y, T)\). We also denote by \( \text{APSO}(Y) \) the class of all algebraic polynomial Stein operators for the target random variable \( Y \).

Proposition 3.2. Fix \( d \geq 1 \). Assume that the target random variable \( Y \) has the Gaussian polynomial form given by relation (3.1). Then, the following statements are in order:

(a) \( \text{APSO}(Y) \subseteq \text{PSO}(Y) \).

(b) When \( d = 1 \), we have \( \text{APSO}(Y) = \text{PSO}(Y) \) (or equivalently \( \text{ASC}(Y, T) = \text{SC}(Y, T) \) for every \( T \geq 1 \)).

Proof. (a) Obvious. (b) Let \( S = \sum_{t=0}^{T} p_t \partial^t \in \text{PSO}(Y) \). Then by definition,

\[
\mathbb{E} [g_T(X) + p_T(Y) | Y] = \mathbb{E} [g_T(X) + p_T(h(X)) | h(X)] = 0. 
\]

Then, Lemma B.1 implies that either \( g_T(X) + p_T(h(X)) = 0 \) (that means \( S \in \text{APSO}(Y) \), and so nothing left to prove) or \( h(X) \) is an even polynomial and \( g_T(X) + p_T(h(X)) \) is an odd polynomial. Next, first note that for \( h \) being even, and any polynomial \( p \in \mathbb{K}[X] \), the composition polynomial \( p(h(x)) \) is even. Next, we show that \( g_T(X) = \sum_{t=0}^{T-1} \Gamma_{Y,T}^{t-1} p_t(h(X)) \) is an even polynomial too. Note that, linearity of the \( \Gamma_Y \) operator allows us to write

\[
\sum_{t=0}^{T-1} \Gamma_{Y,T}^{t-1} p_t(h(X)) = \Gamma_Y \left( \sum_{t=1}^{T} \Gamma_{Y,T}^{t-1} (p_t(h(X))) \right). 
\]

Next, by definition of the \( \Gamma_Y \) operator we can write \( \Gamma_Y(g(X)) = h'(X) \delta^{-1} g(X) \) for every given polynomial \( g \in \mathbb{K}[X] \). Hence, the \( \Gamma_Y \) operator maps an even polynomial \( g(X) \) to a product of two odd polynomials, and \( \Gamma_Y(g(X)) \) is also even since \( h(X) \) is an even polynomial. The above reasoning yields that the polynomial argument inside the \( \Gamma_Y \) operator in the right hand side (3.3), which is given by

\[
r(X) := p_{T-1}(h(X)) + \ldots + \Gamma_{Y,T}^{t-1}(p_0(h(X))),
\]

is even, and \( \Gamma_Y(r(X)) \) is even as well. Therefore, the polynomial \( g_T(X) + p_T(h(X)) \), being both even and odd, is necessarily zero, i.e. \( S \in \text{APSO}(Y) \).

Corollary 3.1. Let \( X \sim N(0,1) \), and \( Y = h(X) \) so that \( h \in \mathbb{K}[X] \) with \( \deg h \geq 2 \) and \( \mathbb{E}[h(X)] = 0 \). Assume \( S = \sum_{t=0}^{T} p_t \partial^t \in \text{PSO}(Y) \) is a polynomial Stein operator for the target \( Y \). Identify \( \mathbb{K}[y] \cong \mathbb{K}[h(x)] \). Then necessarily, \( p_T(y) \in \mathbb{K}[y] \cap \langle h'(x) \rangle \).

Proof. Proposition 3.2, item (b) implies that \( S = \sum_{t=0}^{T} p_t \partial^t \in \text{APSO}(Y) \). Hence \( g_T(X) + p_T(h(X)) = 0 \). On the other hand, note that

\[
p_T(h(X)) = - \left( \Gamma_Y^{T-1}(p_0) + \ldots + \Gamma_Y(p_{T-1}) \right) = -\Gamma_Y \left( \Gamma_{Y,T}^{T-1}(p_0) + \ldots + p_{T-1}(X) \right) 
\]

\[
= -\left( D Y, \delta^{-1} \left( \Gamma_{Y,T}^{T-1}(p_0) + \ldots + p_{T-1}(X) \right) \right)
\]
Remark 3.2. (i) In several important settings (such as when \( d = 1 \)) algebraic polynomial Stein operators, when they exist, are easier to find compared to general polynomial Stein operators. This is due to the fact that one can settle the problem of finding an algebraic polynomial Stein operator in a completely algebraic framework – using the algebraic operators that are introduced in Section 2 – to bypass the difficult problem of computing the conditional expectation in the last stage of the general Stein chain. In fact, hereafter, this is our major interest that at last successfully leads to many new polynomial Stein operators for Gaussian polynomial targets of very complex probabilistic nature, whose derivation is far beyond available techniques from the literature.

(ii) For a Gaussian polynomial target random variable \( Y \) of the form (3.1), Proposition 3.2, item (b), implies that \( APSO(Y) = PSO(Y) \) when \( d = 1 \). However, in general,

\[
APSO(Y) \neq PSO(Y), \quad \text{when } d \geq 2. \quad (3.4)
\]

This phenomenon is discussed in the forthcoming example, in which the significant role of probability enters through the conditional expectation. As will become clear, in higher dimensions, with the target random variable \( Y \) of the form (3.1), the situation \( \mathbb{E}[g(X_1, \ldots, X_d) \mid Y] = p(Y) \) most often happens for some polynomial \( p \in \mathbb{K}[Y] \), although \( g(X_1, \ldots, X_d) \in \mathbb{K}[X_1, \ldots, X_d] \setminus \mathbb{K}[Y] \).

Example 3.1. (a) Case \( d = 1 \): Let \( X \sim N(0, 1) \). Although Proposition 3.2, part (b), confirms that \( APSO(Y) = PSO(Y) \), we would like to present some concrete illustrative examples for some known polynomial Stein operators. (i) When \( Y = aX + b \), where \( a, b \in \mathbb{R} \), it is clear that \( APSO(Y) = PSO(Y) \). (ii) Let \( Y = aX^2 + bX + c \), where \( a, b \neq 0 \) and \( a, b, c \in \mathbb{R} \). Note that \( \mathbb{E}[Y] = a + c \). So, from the Malliavin–Stein perspective, the most desirable choice for the zeroth polynomial is \( p_0(y) = y - a - c \). Straightforward computations give that

\[
g_1(X) = \Gamma_Y(p_0(Y)) = \Gamma_Y(Y - a - c) = \Gamma_Y(Y) = 2a^2H_2(X) + 3abH_1(X) + (2a^2 + b^2),
\]

and \( \mathbb{E}[g_1(X)] = 2a^2 + b^2 \). The requirement \( a, b \neq 0 \) yields that \( g_1(X) \neq p(Y) \) almost surely for any polynomial \( p \in \mathbb{K}[Y] \). Hence, we need to choose appropriately the next polynomial \( p_1(y) \in \mathbb{K}[y] \) and move to the next stage. To do this, note that

\[
\Gamma_Y(g_1(X)) = 4a^2H_2(X) + 8a^2bH_1(X) + (3ab^2 + 4a^3).
\]

We now aim to find a polynomial \( p_1(y) \) that satisfies the requirement \( \mathbb{E}[g_1(X)] = -\mathbb{E}[p_1(Y)] \) and the terminating condition that \( g_2 = \Gamma_Y(g_1 + p_1(Y)) \in H(Y)^\perp \). The
polynomial \( p_1(y) = -4a(y - c) + (2a^2 - b^2) \) meets these requirements. It is easily checked that \( \mathbb{E}[g_1(X)] = -\mathbb{E}[p_1(Y)] \) and we also have that

\[
g_2(X) = \Gamma_Y(g_1(X)) + p_1(Y) = \Gamma_Y(g_1(X)) - 4a\Gamma_Y(Y) = \Gamma_Y(g_1(X)) - 4a\Gamma_Y(p_0(Y)) = \Gamma_Y(g_1(X)) - 4ag_1(X) = -4a^3H_2(X) - 4a^2bH_1(X) - 4a^3 - ab^2 = -4a^3(Y - c) - ab^2.
\]

We now choose \( p_2(y) = 4a^2(y - c) + ab^2 \), and note that \( g_2 + p_2(y) = 0 \). Thus, the sequence \( (p_0(Y), p_1(Y), p_2(Y)) \) is an algebraic Stein chain of length \( T = 2 \) for the target random variable \( Y \). Finally, applying Proposition 3.1 yields the Stein operator

\[
S = (ab^2 + 4a^2(y - c)) \partial^2 + (2a^2 - b^2 - 4a(y - c)) \partial + (y - c - a),
\]

which was achieved in \( \text{(Gaunt, R. E., 2019, Proposition 2.1)} \) via a different approach.

(b) **General Case:** Let \( \{X_k \sim N(0, 1) : k \in \mathbb{N}\} \) be a family of i.i.d. standard Gaussian random variables.

(i) Consider the target random variables \( Y = \sum_{k=1}^{K} \alpha_k H_2(X) \) (living in the second Wiener chaos) where the \( \alpha_k \) are non-zero real numbers. Note that by choosing \( \alpha_k = \alpha \) for \( k = 1, \ldots, K \), the random variable \( Y \) boils down to that of the centered gamma distribution with \( K \) degrees of freedom, and by choosing \( \alpha_k = \alpha, -\beta \) for \( k = 1, \ldots, r \), and \( k = r + 1, \ldots, 2r (=K) \), respectively, where \( \alpha \beta > 0 \), the random variable \( Y \) boils down to that of the centered variance-gamma distribution \( \text{Gaunt, R. E. (2014)} \) with parameters \( (r, \theta, \sigma) \), where \( \theta = \alpha - \beta \) and \( \sigma = 2\sqrt{\alpha \beta} \). In \( \text{Arras, B. et al. (2020)} \), the authors provided the following polynomial Stein operator of order \( T \) with linear coefficients for the target random variable \( Y \):

\[
S = \sum_{k=1}^{T} (b_k - a_k y) \partial^k - a_0 y,
\]

where the constants \( (a_k, b_k : k = 0, \ldots, T) \) are explicit and given in \( \text{(Arras, B. et al., 2020, Section 2.1)} \), and \( T \) is the number of distinct coefficients in \( (a_k, k = 1, \ldots, K) \).

Next, relying on \( \text{(Azmoodeh, E., Peccati, G. and Poly, G., 2015, Lemma 3)} \), relation (23), we can easily infer that \( S \in \text{APSO}(Y) \) is an algebraic polynomial Stein operator for \( Y \). Up to now, only when \( K = 1 \) (equivalently \( d = 1 \), and corresponding to a scaled centered gamma random variable) do we know that \( \text{APSO}(Y) = \text{PSO}(Y) \). For example, with the product Gaussian distribution \( Y = X_1 X_2 \) we do not know whether \( \text{APSO}(Y) = \text{PSO}(Y) \).

(ii) The following example provides a polynomial Stein operator that is not algebraic with \( d = 2 \). Consider the target \( Y = X_1^2 X_2^2 \), a product of two independent chi-square random variables, each with one degree of freedom. In \( \text{Gaunt, R. E. (2018)} \), the
author provided the following second order polynomial Stein operator

\[
S = y^2 \partial_y^2 + 2y \partial_y + \frac{1}{4}(1 - y).
\] (3.6)

Using some straightforward computations, one can see that

\[
g_1(X_1, X_2) = \Gamma_Y(p_0(Y)) = -\frac{1}{4} \Gamma_Y(Y) = -\frac{Y}{4} \left( H_2(X_1) + H_2(X_2) + 1 \right) \in \mathbb{K}[X_1, X_2].
\]

Also, \( g_2(X_1, X_2) \) contains a term in the eighth Wiener chaos of the form \( H_6(X_1)H_2(X_2) \) that does not appear in the chaos expansion of \( Y^2 \). This yields that \( g_2(X_1, X_2) + p_2(Y) = g_2(X_1, X_2) + Y^2 \neq 0 \) almost surely, and hence, \( S \not\in \text{APSO}(Y) \) is not an algebraic polynomial Stein operator for the target \( Y \). It is an interesting problem to determine whether \( \text{APSO}(Y) \neq \emptyset \).

(iii) The following example provides a polynomial Stein operator that is not algebraic with \( d = 3 \). Let \( Y = X_1X_2X_3 \) be the product of three independent standard Gaussian random variables. In Gaunt, R. E. (2017), the author provides the following third order polynomial Stein operator for the target random variable \( Y \):

\[
S = y^2 \partial_y^3 + 3y \partial_y^2 + \partial_y - y \in \text{PSO}(Y).
\]

Some tedious computations yield that

\[
g_1(X_1, X_2, X_3) = -\Gamma_Y(Y)
\]

\[
= -\frac{1}{4} \left( H_2(X_1)H_2(X_2) + H_2(X_2) + H_2(X_3) + 1 \right)
\]

\[
- \frac{1}{3} \left( H_2(X_1)H_2(X_3) + H_2(X_1) + H_2(X_3) + 1 \right)
\]

\[
- \frac{1}{3} \left( H_2(X_1)H_2(X_2) + H_2(X_1) + H_2(X_2) + 1 \right) \in \mathbb{K}[X_1, X_2, X_3],
\]

\[
g_2(X_1, X_2, X_3) = -\Gamma_Y(g_1 + p_1) = -\Gamma_Y(g_1 + 1) = -\Gamma_Y(g_1)
\]

\[
= -\frac{1}{3} Y \left( H_2(X_1) + H_2(X_2) + H_2(X_3) \right) - 2Y \in \mathbb{K}[X_1, X_2, X_3],
\]

and

\[
\Gamma_Y(g_2(X_1, X_2, X_3)) = -\Gamma_Y \left( \Gamma_Y(g_1) \right)
\]

\[
= -\frac{1}{15} \left( H_2(X_1)X_2^2X_3^2 + \frac{1}{3} H_3(X_1)X_1X_3^2 + \frac{1}{3} X_1H_3(X_1)X_2^2 \right)
\]

\[
- \frac{1}{15} \left( H_2(X_2)X_1^2X_3^2 + \frac{1}{3} H_3(X_2)X_2X_3^2 + \frac{1}{3} X_2^2H_3(X_3) \right)
\]

\[
- \frac{1}{15} \left( X_1^2X_2^2H_2(X_3) + \frac{1}{3} X_1^2X_3H_3(X_3) + \frac{1}{3} X_2^2X_3H_3(X_3) \right)
\]

\[
- \frac{1}{4} g_1(X_1, X_2, X_3) \in \mathbb{K}[X_1, X_2, X_3].
\]
Hence, \( g_3(X_1, X_2, X_3) + p_3(Y) = \Gamma_Y (g_2(X_1, X_2, X_3)) - 3g_1(X_1, X_2, X_3) + Y^2 \neq 0 \) almost surely, whilst \( \mathbb{E}[g_3(X_1, X_2, X_3) + p_3(Y) | Y] = 0 \) almost surely. In contrast to the product of two independent standard Gaussian random variables, we do not know yet whether the product of three or more independent standard Gaussian random variables admit any algebraic polynomial Stein operator.

**Remark 3.3.** (i) The following algorithm can be used to produce the associated algebraic Stein chains:

**Algorithm : Producing forward algebraic Stein chain**

(a) **At stage** \( t = 0 \), pick a polynomial \( p_0 \in \mathbb{K}[Y] \) such that \( \mathbb{E}[p_0(Y)] = 0 \), and set \( g_0 = -\mathbb{E}[p_0(Y)] = 0 \).

(b) **At stage** \( t > 1 \), with chosen polynomials \( (p_0, p_1, \ldots, p_{t-1}) \) in the subring \( \mathbb{K}[Y] \) satisfying moment property \( \mathbb{E}[p_s(Y)] = -\mathbb{E}[g_s] \) for \( s = 0, \ldots, t-1 \), if

\[
g_t := \Gamma_Y (g_{t-1} + p_{t-1}(Y)) \in \mathbb{K}[Y]
\]

set \( p_t(Y) = -g_t \) and stop.

(c) Otherwise, choose a polynomial \( p_t(Y) \in \mathbb{K}[Y] \) with \( \mathbb{E}[p_t(Y)] = -\mathbb{E}[g_t] \), and continue to stage \( (t+1) \).

(ii) We have not yet explained how to choose the polynomials \( p_t(Y) \) and how to determine whether algebraic Stein chains of finite length beginning with initial state \( p_0(Y) \) do exist. This is in fact the topic of Section 3.3. The question of how to chose the initial state is also of interest. In the Malliavin-Stein method, it is often desirable to take \( p_0(y) = cy \) as the zeroth-order term; however, other choices of initial state can lead to Stein operators with smaller degree \( m \) or smaller order \( T \) (see Table 1). Remark 3.8 describes how initial states can be chosen to obtain such Stein operators.

(iii) Instead of looking for polynomial coefficients in \( \mathbb{K}[Y] \), we could restrict the coefficients to be in a linear subspace, as for example

\[
\mathbb{L} = \mathbb{K}_r[Y] = \{ p(Y) : \text{polynomials with deg}(p) \leq r \}.
\]

This observation is vital for our final goal of finding an algebraic Stein chain by implementing a MATLAB code (see Remark 3.7, item (ii)). We could also formulate the problem with a sequence of possibly different subspaces \( \mathbb{L}_t \subseteq \mathbb{K}[Y] \) at each stage \( t \in \mathbb{N} \). Note also that if

\[
\langle \nabla h(X_1, \ldots, X_d) \rangle \cap \mathbb{K}_r[Y] = \{0\},
\]

there cannot be any non-trivial finite order algebraic polynomial Stein operator satisfying \( \text{deg} p_t(y) \leq r \), \( \forall t \).
3.2. Backward Stein chain

We may also validate an algebraic Stein chain by starting from the highest order term and checking recursively the lower order terms. We illustrate this subsidiary approach only in the univariate case \( d = 1 \) (see Remark 3.4 for the multidimensional case). We start from the highest order derivative coefficient \( p_T(Y) \in \langle \partial h(X) \rangle \cap \mathbb{K}[Y] \) (if this condition is not satisfied, we stop with a negative answer; there cannot be an algebraic polynomial Stein operator with highest order coefficient \( p_T(Y) \)). Let

\[
q_T(X) = -p_T(h(X))/\partial h(X).
\] (3.7)

After \( t \) successful recursion steps, we should have

\[
\delta q_{T-t}(X) \in \left( \langle \partial h(X) \rangle + \mathbb{K}[Y] \right).
\] (3.8)

If this condition is not satisfied we stop with a negative answer; otherwise, there is a polynomial of the target \( p_{T-t-1}(Y) \) such that \( (\delta q_{T-t}(X) - p_{T-t-1}(h(X))) \in \langle \partial h(X) \rangle \). We set

\[
q_{T-t-1}(X) = (\delta q_{T-t}(X) - p_{T-t-1}(h(X)))/\partial h(X)
\]

and continue. The algebraic backward Stein chain ends successfully after \( (T-1) \) backward steps when

\[
\delta q_T(X) = p_0(h(X)) = p_0(Y) \in \mathbb{K}[Y].
\]

In such case

\[
\mathbb{E}[p_0(Y)f(Y)] = \mathbb{E}[\delta(q_1(X))f(Y)] = \mathbb{E}[q_1(X)\partial h(X)\partial f(Y)]
= \mathbb{E}\left[ \frac{\delta q_2(X) - p_1(Y)}{\partial h(X)} \partial h(X)\partial f(Y) \right] = -\mathbb{E}\left[ p_1(Y)\partial f(Y) \right] + \mathbb{E}\left[ \delta q_2(X)\partial f(Y) \right] = \ldots
= -\mathbb{E}\left[ p_1(Y)\partial f(Y) \right] - \mathbb{E}\left[ p_2(Y)\partial^2 f(Y) \right] - \ldots - \mathbb{E}\left[ p_{T-1}(Y)\partial^{T-1} f(Y) \right]
+ \mathbb{E}\left[ \delta q_T(X)\partial^{T-1} f(Y) \right],
\]

with

\[
\mathbb{E}[\delta q_T(X)\partial^{T-1} f(Y)] = \mathbb{E}[q_T(X)\partial h(Y)\partial^T f(Y)] = -\mathbb{E}[p_T(Y)\partial^T f(Y)],
\]

which means that \( S_f(y) = \sum_{t=0}^{T} p_t(y)\partial^t f(y) \) is an algebraic polynomial Stein operator for the target \( Y = h(X) \).

**Remark 3.4.** In this remark \( X = (X_1, \ldots, X_d) \). The backward construction of an algebraic Stein chain in the multidimensional case \( (d > 1) \) is essentially the same as in the univariate case \( d = 1 \) with only the minor difference that at each backward step the divisor polynomials \( q_{T-t} \), \( t = 0, \ldots, T-1 \), are not unique, unlike the univariate case.
For example, start with the polynomial coefficient \( p_T(Y) \) of the highest order term that must be in \( \mathbb{K}[Y] \cap (Dh(X)) \). (If this condition is not valid, we know there is no algebraic polynomial Stein operator with the highest order coefficient term \( p_T \). This condition can be checked by the multivariate polynomial division algorithm, using Gröbner basis.) The latter means there is some polynomial vector \( q_T(X) = (q_{T,1}(X), \ldots, q_{T,d}(X)) \in \mathbb{K}[X]^d \) such that

\[
p_T(h(X)) = \sum_{k=1}^{d} q_{T,k}(X) D_k h(X).
\]

However, when the condition is satisfied, (when \( d > 1 \)) the divisor polynomial \( q_T(X) \) does not need to be unique. In fact, one can add any solution of the homogeneous equation

\[
\sum_{k=1}^{d} u_{T,k}(X) D_k h(X) = 0.
\]

This point has to be compared with relation (3.7) in the univariate case \( d = 1 \) where the divisor polynomial \( q_T \) is unique.

### 3.3. Null controllability

This section outlines the connection between Stein’s method and the theory of linear systems. For a comprehensive account of the theory of linear systems, the reader is referred to Rugh, W. (1993); Fuhrmann, P. A. (1981); Fuhrmann, P. A. and Helmke, U. (2015); Bourlès, H. (2010); Bourlès, H. and Marinescu, B. (2011) and references therein. Let \( \mathbb{V} \) (state space) and \( \mathbb{L} \) (control space) be two vector spaces (not necessarily finite dimensional) over \( \mathbb{K} \). A linear discrete system \( \Sigma \) is a quadruple \( (\mathbb{V}, \mathbb{L}, \Gamma, \Lambda) \), where \( \Gamma : \mathbb{V} \to \mathbb{V} \) (evolution operator) and \( \Lambda : \mathbb{L} \to \mathbb{V} \) (input operator) are two linear maps, and the state variables follow the following linear dynamic

\[
g_t = \Gamma g_{t-1} + \Lambda p_{t-1}, \quad t = 1, 2, \ldots
\]

Let us introduce the notion of (null) controllability in the theory of linear systems, another component that plays a significant role in our paper.

**Definition 3.3.** Let \( T \geq 1 \). We say that an initial state \( g_0 \in \mathbb{V} \) is null controllable in \( T \)-steps and denoted by \( g_1 \xrightarrow{\mathbb{V}} 0 \), if there is a finite (null) control sequence \( \{p_0, \ldots, p_{T-1}\} \subseteq \mathbb{L} \) such that the linear recursion (3.9) reaches \( g_T = 0 \in \mathbb{V} \) at time \( T \). Denote

\[
\mathcal{C}(\Sigma, T) = \left\{ g_0 \in \mathbb{V} : g_0 \xrightarrow{\mathbb{V}} 0 \right\}.
\]

The set of all null controllable states is \( \mathcal{C}(\Sigma) = \bigcup_{T \geq 1} \mathcal{C}(\Sigma, T) \). Clearly, both \( \mathcal{C}(\Sigma, T) \) and \( \mathcal{C}(\Sigma) \) are vector spaces over \( \mathbb{K} \), and \( \mathcal{C}(\Sigma, T) \subseteq \mathcal{C}(\Sigma, T+1) \) for every \( T \geq 1 \).
Lemma 3.1. Consider a linear discrete system $\Sigma$ as described above. Then, the state $g_0 \in V$ is null controllable if and only if for some $T \geq 1$,

$$\Gamma^T g_0 \in \sum_{t=1}^{T} \Gamma^{T-t} \Lambda(L).$$

Proof. Using the linear dynamic (3.9) one can readily obtain that, for $t = 1, \ldots, T$,

$$g_t = \Gamma^t g_0 + \sum_{s=1}^{t} \Gamma^{t-s} \Lambda p_{s-1},$$

and hence, the result follows at once.

Remark 3.5. If the null controllability problem has finite horizon solutions the following algorithm finds the shortest null control sequences:

\begin{algorithm}
At stage $t = 0$, pick an initial state $g_0 \in V$, and consider the linear equation

$$\Gamma g_0 = -\Lambda p_0.$$ 

If this equation has a solution in $L$, then stop (equivalently, $\Gamma(-g_0) \in \text{Range}(\Lambda)$).

Otherwise we continue the recursion, until at some stage $T$ for the first time the system

$$\Gamma^T g_0 = -(\Lambda p_{T-1} + \Gamma \Lambda p_{T-2} + \Gamma^2 p_{T-3} + \cdots + \Gamma^{T-2} \Lambda p_1 + \Gamma^{T-1} \Lambda p_0)$$  (3.10)

has solution (which is not necessarily unique) $(p_0, \ldots, p_{T-1}) \in L^T$.
\end{algorithm}

The linear control framework applies directly to our forward algebraic Stein chain construction. To see this, consider the linear system $\Sigma$, where the state space $V = K[X_1, \ldots, X_d]$ and the control space $L = K[Y]$. We stress that both the state and control spaces are infinite dimensional vector spaces which do not enjoy suitable analytic properties such as being a Banach space. Moreover, the evolution and input operators are given by the Malliavin operator $(X = (X_1, \ldots, X_d))$,

$$\Gamma_Y : K[X_1, \ldots, X_d] \to \left\langle \partial X_1 h(X), \ldots, \partial X_d h(X) \right\rangle \subseteq K[X_1, \ldots, X_d].$$

The following result is an immediate consequence of the definitions of an algebraic Stein chain and a null controllable state.

Proposition 3.3. Fix $T \geq 1$. Let $d \geq 1$, and assume that the Gaussian polynomial target variable $Y$ takes the form (3.1). Let $\Sigma = (V, L, \Gamma_Y, \Gamma_Y)$ be the linear system as described above. Let

$$K_0[Y] = \left\{ p(Y) \in K[Y] : E[p(Y)] = 0 \right\}.$$  (3.11)
Remark 3.6. Assume the convention $X = (X_1, \ldots, X_d)$ as above. Let $p_0 \in \mathbb{L} = \mathbb{K}[Y]$.
Then, the target random variable $Y = h(X)$ admits an algebraic polynomial Stein operator with zeroth order term $p_0$ if the initial state $g_0 = g_0(X) : = \Gamma_Y p_0(h(X))$ can be null controllable in a finite time. Mathematically, for some $t \geq 1$, there exist polynomials $p_0(Y), \ldots, p_{t-1}(Y) \in \mathbb{K}[Y]$ such that
\[
g_{t+1}(X) = \Gamma^t g_0(X) + \sum_{s=1}^{t} \Gamma_Y^{t-s} p_{s-1}(h(X)) = 0.
\] (3.12)

The latter is a direct consequence of the set inclusion $\Gamma_Y^t(\mathbb{K}[Y]) \subseteq \sum_{s=1}^{t} \Gamma_Y^{t-s}(\mathbb{K}[Y])$ which is in order as soon as the ascending chain $I_1 \subseteq \cdots \subseteq I_t \subseteq I_{t+1} \subseteq \cdots$ would stop, where $I_t = \sum_{s=1}^{t} \Gamma_Y^{t-s}(\mathbb{K}[Y])$ for $t \in \mathbb{N}$. Although, the Hilbert base theorem guarantees that $\mathbb{K}[X]$ is a Noetherian ring, however we cannot conclude that the chain would necessary stop due to the undesirable fact that the vector spaces $I_t$ are not ideal.

Remark 3.7. (i) In order to apply the rich theory of linear control systems in our framework, we mention the following issues. Firstly, in our formulation, both the state and the control spaces are infinite dimensional vector spaces over $\mathbb{K}$. Secondly, it is also possible that one can think of the (infinite dimensional) Hilbert state space $\mathbb{V} = L^2(\mu)$ with $\mu$ the standard Gaussian measure on the real line. However, the control space $\mathbb{L}$ “must” be taken as the ring $\mathbb{K}[Y]$, which is not a Hilbert space (not even a Banach space). Therefore, we believe that the theory developed for infinite dimensional linear systems, see for instance Fuhrmann, P. A. (1981); Triggiani, R. (1975, 1976), is hardly applicable in our setting to study the null controllability.

(ii) To bypass the obstacles mentioned above, in the computer implementation of the algorithm we may set an upper bound on the polynomial degree of the Stein operator coefficients, assuming that the initial state and controls belong to $\mathbb{L} = \mathbb{K}_m[Y]$. Then, at stage $t$,
\[
\deg(g_t) \leq \begin{cases} 
\deg(h) \times m + (\deg(h) - 2) \times t & \text{(univariate $h$)} \\
\deg(h) \times (m + t) & \text{(multivariate $h$)}
\end{cases}
\]
Hence, in our linear system, state space is time varying (increasing in time); however, within a finite time horizon $T$ we can implement the algorithm to check null controllability up to time $T$ on a finite dimensional state space $\mathbb{R}^{\text{deg}(h) \times m + (\text{deg}(h) - 2) \times T}$ and control space $\mathbb{R}^{\text{deg}(h) \times m}$, see also Remark 3.3. Moreover, we point out that even by fixing a time horizon $T$, the Kalman criterion as described in the introduction cannot be used to study null controllability, because the criterion checks null controllability within the whole null states $\mathcal{C}(\Sigma)$ and not on $\mathcal{C}(\Sigma, T)$. Lastly, in the computer implementation of the linear system $\Sigma$ (as described above) the input operator $\Lambda = \Gamma_Y \Theta$, where the operator $\Theta$ is given by the embedding $p(Y) \in \mathbb{K}[Y]$ into $p(h(X_1, \ldots, X_d)) \in \mathbb{K}[X_1, \ldots, X_d]$. This is due to the fact that operator $\delta^{-1}$ (in the definition of the $\Gamma_Y$ operator) acts over the polynomial ring $\mathbb{K}[X_1, \ldots, X_d]$.

**Remark 3.8.** In the Malliavin-Stein method, the zeroth-order polynomial $p_0(y) = cy$ is commonly used. However, other choices of $p_0(y)$ may allow for Stein operators with either lower order $T$ or lower maximal coefficient degree $m$. In the computer implementation of our algorithm, for a given target $Y = h(X)$ the input is a zeroth-order polynomial coefficient $p_0(h(x))$, with $E[p_0(h(X))] = 0$. In order to find a Stein operator with generic $p_0(y)$ we run the algorithm several times with the respective initial coefficients $p_0^{(k)}(y) = y^k - E[h(X)^k]$, in order to obtain the corresponding Stein operators satisfying

$$E[S^{(k)}f(Y)] = -E[p_0(Y)f(Y)]$$

for $k = 1, \ldots, m_0$. When the Stein chains end successfully and Stein operators are found for the initial coefficients above, by solving a linear system for the linear combination one can further reduce the order or the maximal coefficient degree of a Stein operator with a generic zeroth-order polynomial coefficient $p_0(y)$ of degree $\leq m_0$.

4. Applications to Gaussian Hermite polynomials

4.1. Highest order polynomial coefficient

Before the next proposition, we note the following lemma that will be needed in the proof. The result of Indritz, J. (1961) is stated for the physicists' Hermite polynomials rather than the probabilists’ Hermite polynomials, as used in our paper.

**Lemma 4.1.** (Indritz, J. (1961)). Define $E_p(x) = (\pi^{1/2}2^{2p}p!)^{-1/2}e^{-x^2/4}H_p(x)$. Then the relative maxima of $|E_p(x)|$, $x \geq 0$, steadily increase, i.e., if $x_1 < x_2 < \ldots < x_j$ are the non-negative zeros of $E'_p$ for fixed $p$, then

$$|E_p(x_1)| < |E_p(x_2)| < \ldots < |E_p(x_j)|.$$

**Proposition 4.1.** Let $X \sim N(0, 1)$. Assume that $Y = H_p(X)$, where $H_p$ is the Hermite polynomial of degree $p \geq 2$ (the case $p = 1$ corresponds to standard Gaussian distribution
that is not of interest in this paper). Let \( S = \sum_{i=0}^{T} p_i \partial^i \in PSO(Y) = APSO(Y) \) be a polynomial Stein operator for \( Y \). As before, identify the (dependent) variable \( y = H_p(x) \). Then the following properties hold:

(a) \( \mathbb{K}[y] \cap (H_p'(x)) \) is an ideal of the subring \( \mathbb{K}[y] \). Moreover,
\[
\mathbb{K}[y] \cap (H_p'(x)) = \langle t(y) \rangle
\]
with
\[
t(y) = \prod_{z : H_p'(z) = 0} (y - H_p(z)).
\]

(b) \( p_T(y) \in \mathbb{K}[y] \cap (H_p'(x)) \), and \( p_T(H_p(x)) = 0 \) for all solutions \( x \) of \( H_p'(x) = 0 \). In particular, \( p_T(y) = 0 \) at all local minimum or local maximum values \( y \) of \( H_p \).

(c) \( \deg(p_T) \geq p/2 \) if \( p \) is even, and \( \deg(p_T) \geq p - 1 \) if \( p \) is odd.

**Proof.** (a) It is easy to see that \( \mathbb{K}[y] \cap (H_p'(x)) \subseteq \mathbb{K}[y] \) is an ideal of the subring \( \mathbb{K}[y] \). Also, after substitution \( y = H_p(x) \), it becomes clear that the polynomial \( t(y) = t(H_p(x)) \), given by \( t(y) = \prod_{z : H_p'(z) = 0} (y - H_p(z)) \in \mathbb{K}[y] \), is divisible by \( H_p'(x) \) by the Taylor expansion \( H_p(x) - H_p(z) = \sum_{k=1}^{p} H_p^{(k)}(z)(x-z)^k \) for every \( z \) such that \( H_p'(z) = 0 \). Note that \( H_p'(x) = pH_{p-1}(x) \) and all the roots of the Hermite polynomials are real. So the claim follows at once by a direct application of Remark 2.1.

(b) Apply Corollary 3.1. The rest are direct consequences.

(c) This follows from part (b) and the fact that \( H_p \) has \( p/2 \) distinct values for the local maxima and minima when \( p \) is even, and \( p - 1 \) distinct values for the local maxima and minima when \( p \) is odd. It is a standard property of \( H_p \) that it has exactly \( p \) real roots. Therefore \( H_p \) must have \( p \) local maxima and minima when \( p \) is even and \( p - 1 \) local maxima and minima when \( p \) is odd. But when \( p \) is even, \( H_p \) is an even function and there can hence be at most \( p/2 \) distinct values for the local maxima and minima. That there are exactly \( p/2 \) distinct local maxima and minima in the even case now follows from the stronger result of Lemma 4.1. For odd \( p \), we let \( x_1 < x_2 < \ldots < x_{(p-1)/2} \) be the non-negative zeros of \( H_p' \), and as \( H_p \) is an odd function we have that \( -x_{(p-1)/2} < \ldots < -x_2 < -x_1 \) are the negative zeros of \( H_p' \). Then due to the stronger result of Lemma 4.1, we have that \( |H_p(x_1)| = |H_p(-x_1)| < \ldots < |H_p(x_{(p-1)/2})| = |H_p(-x_{(p-1)/2})| \) with \( H_p(x_k) = -H_p(-x_k) \) for all \( k = 1, 2, \ldots, (p-1)/2 \). Thus, there are exactly \( p - 1 \) distinct values for the local maxima and minima in the odd case.

**Remark 4.1.** It is well-known that (see, (Gaunt, R. E., 2019, Proposition 2.1)) target random variables \( Y = h(X) \) with \( h \in \mathbb{K}[X] \) and \( \deg(h) = 2 \) admit a polynomial Stein operator of order two with linear polynomial coefficients.

**Remark 4.2.** Part (c) of Proposition (4.1) is useful in implementing our code to find Stein operators for \( H_p(X) \). In particular, if \( p \) is even we must seek polynomial Stein...
operators with coefficients that have degree at least \( p/2 \), and if \( p \) is odd we require the degree to be at least \( p - 1 \). As can be seen in Table 1, we have used our MATLAB code to find polynomial Stein operators for \( H_p \), \( p = 1, \ldots, 10 \), that attain these minimum possible degrees. In fact, we have also tested this for \( p = 11, 12, 14, 16, 18 \), and our MATLAB code has always been able to find a Stein operator with the minimum possible degree. Therefore, it seems reasonable to conjecture that this is the case for all \( p \geq 1 \).

4.2. Examples of Stein operators for \( H_p(X) \)

For univariate Gaussian polynomials, our MATLAB code can find all polynomial Stein operators up to a given order \( T \) and maximum degree \( m \). We illustrate this in Appendix A by providing examples of Stein operators for \( H_p(X) \), \( p = 1, \ldots, 6 \), where \( X \sim N(0,1) \). The code can also be applied to \( h(X) \) when \( h \) is not a Hermite polynomial, as demonstrated in the Supplementary Information (SI).

In this paper, we only give the ‘simplest’ Stein operators for \( H_p(X) \), \( p = 1, \ldots, 6 \). For \( p = 1, \ldots, 6 \), we list the Stein operators with the lowest order \( T \) and lowest degree \( m \) (if there are two Stein operators with the lowest \( T \) and lowest \( m \)), we give the one with lowest \( m \) (\( T \)). For \( p \geq 7 \), the Stein operators become too complex to state in this paper, but in Table 1 we give a summary of the complexity of the ‘simplest’ Stein operators for \( p = 1, \ldots, 10 \). Many other examples are given in the SI, some of which are ‘simpler’ in other senses than those presented in Appendix A, such as having lower values of \( T + m \) or \( T \times m \).

The Stein operators in Table 1, except the one for \( H_9(X) \) with \((T, m) = (9, 40)\), were obtained using a standard laptop, a MacBook Pro with processor: 2.9 GHz Dual-Core Intel Core i5 and memory: 8 GB, 2133 MHz, LPDDR3. Despite the complexity of the operators, the code found them rather quickly. For example, it was able to find the horribly complex Stein operator for \( H_9(X) \) with \((T, m) = (41, 8)\) in just 343.3 s. We easily obtained the Stein operator for \( H_9(X) \) with \((T, m) = (9, 40)\) using a powerful computer server at the University of Manchester, and, whilst not explored in this study, a full exploitation of such computational power could yield some hugely complex Stein operators! Indeed, using

| Distribution | General 0th-order term | 0th-order term cyf(y) Min T | Min m | CPU | Elapsed time |
|--------------|------------------------|-----------------------------|-------|-----|-------------|
| \( H_1(X) \) | (1, 1) | (1, 1) | (1, 1) | 0.4 | 0.5 s |
| \( H_2(X) \) | (1, 1) | (1, 1) | (1, 1) | 0.4 | 0.5 s |
| \( H_3(X) \) | (3, 4) | (5, 2) | (4, 3) | 0.7 | 0.9 s |
| \( H_4(X) \) | (2, 3) | (3, 2) | (3, 2) | 0.6 | 0.8 s |
| \( H_5(X) \) | (5, 12) | (13, 4) | (6, 11) | 1.4 | 4.6 s |
| \( H_6(X) \) | (3, 6) | (6, 3) | (4, 5) | 1.3 | 2.1 s |
| \( H_7(X) \) | (7, 24) | (25, 6) | (8, 23) | 5.8 | 44.4 s |
| \( H_8(X) \) | (4, 10) | (10, 4) | (5, 9) | 1.8 | 5.2 s |
| \( H_9(X) \) | (9, 40) | (41, 8) | (10, 39) | 21.0 | 343.3 s |
| \( H_{10}(X) \) | (5, 15) | (15, 5) | (6, 14) | 3.6 | 21.8 s |
just a standard PC, we obtained a Stein operator for $H_{20}(X)$ with zero-order term $cyf(y)$ and $(T, m) = (11, 54)$, with 402.6 CPU and elapsed time 4276 s.

There are several interesting observations we can make from the Stein operators presented in Appendix A and the summary in Table 1. We notice that there is an important increase in complexity from $p = 1, 2$ to $p = 3, 4$, in which the Stein operators go from being first order with linear coefficients (for which it is simple to solve the corresponding Stein equation) to being at least second order with higher order coefficients (for which we are not able to solve the corresponding Stein equation). There is a further increase in complexity from $p = 3, 4$ to $p = 5, 6$, in which the Stein operators go from being expressed in one line equations to equations that sprawl several lines.

From Table 1, we observe that, for 0-order term $cyf(y)$ the ordered pairs $(T, m)$ satisfy the following recipes:

(a) Minimum $T$: (i) when $p \geq 4$ is even, we have $(T, m) = (p/2 + 1, \binom{p}{2} + p/2 - 1)$; (ii) when $p \geq 3$ is odd, we have $(T, m) = (p + 1, \binom{p}{2} + (p - 1)/2 - 1)$.

(b) Minimum $m$: (i) when $p \geq 2$ is even, we have $(T, m) = (p/2 + 1, \binom{p}{2} + 1)$; (ii) when $p \geq 3$ is odd, we have $(T, m) = (p - 2)/2 - (p - 1)/2, p - 1)$.

For general 0-order term we have:

(c) Minimum $T$: (i) when $p$ is even, just pick up the pair $(T, m)$ minimizing $m$ with 0-order term $cyf(y)$ associated with that value of $p$ and switch the components; (ii) when $p$ is odd, pick up the pair $(T, m)$ minimizing $T$ with 0-order term $cyf(y)$ associated with that value of $p$ and set $(T - 1, m + 1)$.

(d) Minimum $m$: for both even and odd $p$ this is the same as in item (b).

Rather curiously, the minimum values of $T$ seem to be connected with the number of distinct local maxima and minima of the Hermite polynomial $H_p$. We know that this is the case for the minimum possible degree $m$, due to Proposition 4.1, part (c).

We observe that for each $p$ we tested the code always found a Stein operator with the minimum possible degree (see Proposition 4.1, part (c)). We do not, however, have a proof of an analogous result for the minimum possible order $T$ of Stein operators. In theory, it is therefore possible that we have not actually found Stein operators with the minimum possible values of $T$. However, our tests suggest that this is a remote possibility. For example, for zero-order term $cyf(y)$, our code could not find any Stein operators for $H_{4}(X)$ with input variables $T = 2$ and $m = 60$, nor for $H_{5}(X)$ with input variables $T = 5$ and $m = 80$. Additionally, we performed many other tests with our code and verified all of the predicted recipes (a) – (d) for a number of values of $p$ between 11 and 20 (the complex Stein operator for $H_{12}(X)$ is given in the SI). As such, we believe that it is reasonable to conjecture that the recipes (a) – (d) hold for all $p$. 
Appendix A: Stein operators for univariate Gaussian Hermite polynomials

The Stein operator for $H_1(X)$ is the classical standard Gaussian Stein operator of Stein, C. (1972). The Stein operator for $H_2(X)$ is a special case of the Stein operator for $aX^2 + bX + c$, $a, b, c \in \mathbb{R}$, of Gaunt, R. E. (2019). Also, the Stein operators (A.1) and (A.2) were already obtained by Gaunt, R. E. (2019). All other Stein operators in this appendix are new. Many more examples are given in the Supplementary Information.

$H_1(X)$ and $H_2(X)$:

$$y - \partial, \; y - (2y + 2)\partial$$

$H_3(X)$:

$$y - 6\partial - 99y\partial^2 + (216 - 27y^2)\partial^3 + 486y\partial^4 + (486y^2 - 1944)\partial^5 \quad (A.1)$$

$$(290y - y^3) + (528y^2 - 1560)\partial + (243y^3 - 1404y)\partial^2 + (27y^4 - 648y^2 + 2160)\partial^3$$

$H_4(X)$:

$$y - (24 + 44y)\partial + (576 + 144y - 16y^2)\partial^2 + (192y^2 + 576y - 3456)\partial^3 \quad (A.2)$$

$$(-y^2 + 50y + 24) + (64y^2 + 72y - 1008)\partial + (16y^3 - 48y^2 - 576y + 1728)\partial^2$$

$H_5(X)$:

$$y - 120\partial - 75325y\partial^2 + (-81875y^2 + 7704000)\partial^3 + (-31250y^3 + 270600000y)\partial^4$$

$$+ (-3125y^4 + 527800000y^2 - 39086400000)\partial^5 + (280000000y^3 - 155065000000y)\partial^6$$

$$+ (350000000y^4 - 241335000000y^2 + 14306880000000)\partial^7$$

$$+ (-198750000000y^3 + 5340360000000000y)\partial^8$$

$$+ (-331250000000y^4 + 3495000000000000y^2 - 1170432000000000000y)\partial^9$$

$$+ (390000000000000y^3 - 108432000000000000y)\partial^{10}$$

$$+ (97500000000000000y^4 - 669600000000000000y^2 + 352512000000000000000)\partial^{11}$$

$$+ (-21600000000000000000y^3 + 622080000000000000000y)\partial^{12}$$

$$+ (-1080000000000000000000y^4 + 622080000000000000000000000000y^2 - 2985984000000000000000000)\partial^{13}$$
\( y^9 \) \(- 104800744 \, y^7 + 174104044032 \, y^5 \) \(- 82431615212544 \, y^3 + 9617056740900864 \, y \) \
\(+ (-83053520 y^8 + 191761742080 y^6 - 148596701936640 y^4 + 33440484399022080 y^2 \
\quad - 868706901405204800) \, \partial \
\)+ (-23029125 y^9 + 72332912000 y^7 - 88767223008000 y^5 + 32039796049920000 y^3 \
\quad - 198459365090184000 y) \, \partial^2 \
\)+ (-2831875 y^{10} + 11857320000 y^8 - 22211556000000 y^6 + 11983543971840000 y^4 \
\quad - 1826589574103040000 y^2 + 54875902433034240000) \, \partial^3 \
\)+ (-156250 y^{11} + 855800000 y^9 - 2353387200000 y^7 + 1868056934400000 y^5 \
\quad - 530407371571200000 y^3 + 36302379968102400000) \, \partial^4 \
\)+ (-3125 y^{12} + 22000000 y^10 - 85519200000 y^8 + 99156326400000 y^6 \
\quad - 650653212672000000 y^4 + 19243712957644800000 y^2 - 849260402284953600000) \, \partial^5 \\
H_6(X) : \\
y + (-1278 y - 720) \, \partial + (-972 y^2 + 103230 y + 756000) \, \partial^2 \
\quad + (-216 y^3 + 228960 y^2 + 16491600 y - 120528000) \, \partial^3 \
\quad + (71280 y^3 + 6771600 y^2 - 30715200 y - 3265920000) \, \partial^4 \
\quad + (-314928000 y^2 - 19945440000 y + 12597120000) \, \partial^5 \
\quad + (-209952000 y^3 - 19945440000 y^2 + 251942400000 y + 7558272000000) \, \partial^6 \\
(15303970800 y - 252586320 y^2 - 6227803 y^3 + 599 y^4) \
\quad + (-6722792640000 - 2872324800 y + 30858084000 y^2 - 247410960 y^3 - 6390132 y^4) \, \partial \
\quad + (-251522496000000 - 831421584000 y + 291114000000 y^2 + 141578442000 y^3 \
\quad - 43020180 y^4 - 1746684 y^5) \, \partial^2 \
\quad + (1173771648000000 + 2794694000000 y - 3912572160000 y^2 - 13197168000 y^3 \
\quad + 1633473000 y^4 - 129384 y^6) \, \partial^3 \\

**Appendix B: Auxiliary lemma**

**Lemma B.1.** Fix \( d = 1 \). Let \( X \sim N(0, 1) \). Assume that \( g(X), h(X) \in K[X] \) with \( K = \mathbb{Q}, \mathbb{Z} \). Then with \( Y = h(X) \), 
\[
E[g(X) \mid h(X) = y] = 0, \quad \forall y \in \text{supp}(Y) 
\]  
if and only if one of the two conditions below are satisfied:

(a) \( g \equiv 0 \), i.e., \( g \) is identically the zero polynomial.
(b) $g$ is an odd polynomial and $h$ is an even polynomial, meaning that $g(-x) = -g(x)$ and $h(x) = h(-x)$ for all $x \in \mathbb{R}$.

**Proof.** Note that (B.1) is equivalent to

$$
\sum_{x \in \mathbb{R} : h(x) = y} g(x)e^{-x^2/2} = 0, \quad \forall \ y \in \text{supp}(Y).
$$

(B.2)

Next we discuss on the degree of polynomial $h$. Suppose first that $h$ has odd degree, in which case $h$ is not an even polynomial. Then for all $y$ with $|y|$ large enough the algebraic equation $h(x) = y$ has one and only one real root, let us say $r(y)$, and moreover the mapping $y \mapsto r(y)$ is strictly monotone outside a compact interval. Then relation (B.2) yields that the polynomial $g$ has infinitely many real roots, which implies that $g \equiv 0$.

Now, consider the case with $h$ of even degree. Then, for all $|y|$ large enough, depending on the signs of $y$ and of the leading coefficient of $h$ either equation $h(x) = y$ has no real roots or it has only two real roots, let us say $r_1(y), r_2(y)$ so that $r_1(y) < 0 < r_2(y)$. Moreover, as before the mappings $y \mapsto r_1(y), y \mapsto r_2(y)$ are strictly monotone outside a compact interval. Next assume that $g \not\equiv 0$. In the latter case relation (B.2) means that for all large enough $y$:

$$
g(r_1(y)) \quad \frac{g(r_1(y))}{g(r_2(y))} = -\exp \left( \frac{(r_2(y))^2 - (r_1(y))^2}{2} \right).
$$

(B.3)

Now, by choosing $y \in \mathbb{Q}$, the roots $r_j(y), j = 1, 2$, of the algebraic equations $h(x) = y$, are algebraic numbers (because $\mathbb{K} = \mathbb{Z}, \mathbb{Q}$), and hence the left-hand side of relation (B.3) is an algebraic number since the quotient of two algebraic numbers is again an algebraic number. This implies that the right hand side of relation (B.3) is an algebraic number and that in virtue of the Lindemann-Weierstrass theorem (Baker, A., 1975, Theorem 1.4) this only happens when $r_2(y) = \pm r_1(y)$. But our sign justification yields that must be $r_1(y) = -r_2(y)$ for all large enough $y \in \mathbb{Q}$. Therefore, from the relation (B.3) one can infer that $g(r_2(y)) = -g(r_1(y))$ for all large enough $y \in \mathbb{Q}$. Now consider polynomial $G(x) := g(x) + g(-x)$. Then, this polynomial would have infinitely many roots, and hence $G(x) = 0$ for all $x \in \mathbb{R}$, i.e., $g$ is an odd polynomial. Looking to a new polynomial $H(x) := h(x) - y$, and with a similar argument one can conclude that $h(x) = h(-x)$ for all $x \in \mathbb{R}$, i.e., $h$ is an even polynomial. \qed

**Acknowledgements**

We would like to thank Ivan Nourdin for first bringing to our attention the fascinating problem of finding Stein operators for $H_n(X), n \geq 3$. Without these initial conversations, this paper would not exist. EA would also like to thank Peter Eichelsbacher and Yacine Barhoumi-Andréani for many stimulating discussions on Stein’s method. We also would like to thank our enthusiastic readers Kaie Kubjas and Luca Sodomaco for their comments and remarks. RG is supported by a Dame Kathleen Ollerenshaw Research Fellowship.
References

Antsaklis, P. J. and Michel, A. N. (2007) *A linear systems primer*. Birkhäuser Boston, Inc., Boston, MA.

Arras, B., Azmoodeh, E., Poly, G. and Swan, Y. (2019) A bound on the 2-Wasserstein distance between linear combinations of independent random variables. *Stoch. Proc. Appl.*, **129**, 2341–2375.

Arras, B., Azmoodeh, E., Poly, G. and Swan, Y. (2020) Stein characterizations for linear combinations of gamma random variables. *Braz. J. Probab. Stat.*, **34**, 394–413.

Arras, B. and Houdré, C. (2019) *On Stein's method for infinitely divisible laws with finite first moment*. Springer Briefs.

Arras, A., Mijoule, G., Poly, G. and Swan, Y. (2017) A new approach to the Stein-Tikhomirov method: with applications to the second Wiener chaos and Dickman convergence. https://arxiv.org/abs/1605.06819.

Azmoodeh, E. and Gasbarra, D. (2018) New moments criteria for convergence towards normal product/tetila laws. https://arxiv.org/abs/1708.07681.

Azmoodeh, E., Peccati, G. and Yang, X. (2021) Malliavin-Stein Method: a Survey of Recent Developments. https://arxiv.org/abs/1809.01912.

Baker, A. (1975) *Transcendental Number Theory*. Cambridge University Press, First edition.

Barbour, A. D. (1990) Stein’s method for diffusion approximations. *Probab. Theory Rel.*, **84**, 297–322.

Bourlès, H. (2010) *Linear Systems*. Wiley-series : Control Systems, Robotics and Manufacturing; 1 edition.

Bourlès, H. and Marinescu, B. (2011) *Linear time-varying systems. Algebraic-analytic approach*. Lecture Notes in Control and Information Sciences, 410. Springer-Verlag, Berlin.

Chen, L. H. Y., Goldstein, L. and Shao, Q.–M. (2011) *Normal Approximation by Stein’s Method*. Springer.

Döbler, C. (2015) Stein’s method of exchangeable pairs for the beta distribution and generalizations. *Electron. J. Probab.*, **20**, no. 109, 1–34.

Döbler, C., Gaunt, R. E. and Vollmer, S. J. (2017) An iterative technique for bounding derivatives of solutions of Stein equations. *Electron. J. Probab.*, **22**, no. 96, 1–39.

Döbler, C. and Peccati, G. (2018) The Gamma Stein equation and non-central de Jong theorems. *Bernoulli*, **24**, 3384–3421.

Eden, R. and Viens, F. (2013) General upper and lower tail estimates using Malliavin calculus and Stein’s equations. In *Seminar on Stochastic Analysis, Random Fields and Applications VII* (R. C. Dalang, M. Dozzi and F. Russo, eds.) *Progress in Probability* 67 Birkhäuser/Springer, Basel, 55–84.

Eden, R. and Vquez, J. (2015) Nourdin-Peccati analysis on Wiener and Wiener-Poisson space for general distributions. *Stoch. Proc. Appl.*, **125**, 182–216.
Eichelsbacher, P. and Thäle, C. (2014) Malliavin-Stein method for Variance-Gamma approximation on Wiener space. *Electron. J. Probab.*, **20**, no. 123, 1–28.

Fuhrmann, P. A. (1981) *Linear systems and operators in Hilbert space*. McGraw-Hill International Book Co., New York.

Fuhrmann, P. A. and Helmke, U. (2015) *The Mathematics of Networks of Linear Systems*. Universitext Springer International Publishing.

Gaunt, R. E. (2014) Variance-Gamma approximation via Stein’s Method. *Electron. J. Probab.*, **19**, no. 38, 1–33.

Gaunt, R. E. (2017) On Stein’s method for products of normal random variables and zero bias couplings. *Bernoulli*, **23**, 3311–3345.

Gaunt, R. E. (2018) Products of normal, beta and gamma random variables: Stein operators and distributional theory. *Braz. J. Probab. Stat.*, **32**, 437–466.

Gaunt, R. E. (2019) Stein operators for variables form the third and fourth Wiener chaoses. *Stat. Probabil. Lett.*, **145**, 118–126.

Gaunt, R. E., Mijoule, G. and Swan, Y. (2019) An algebra of Stein operators. *J. Math. Anal. Appl.*, **469**, 260–279.

Gaunt, R. E., Mijoule, G. and Swan, Y. (2020) Some new Stein operators for product distributions. *Braz. J. Probab. Stat.*, **34**, 795–808.

Götze, F. (1991) On the rate of convergence in the multivariate CLT. *Ann. Probab.*, **19**, 724–739.

Indritz, J. (1961) An Inequality for Hermite Polynomials. *P. Am. Math. Soc.*, **12**, 981–983.

Kailath, T. (1980) *Linear Systems*. Prentice-Hall, Englewood Cliffs, NJ.

Kalman, R. E. (1960) On the general theory of control systems. In: *Proceedings of the First World Congress of the International Federation of Automatic Control, 1960*, 481–493.

Kalman, R. E. (1963) Mathematical Description of Linear Dynamical Systems. *SIAM J. Control.*, **1**, 152–192.

Kawano, Y. and Ohtsuka, T. (2016) Commutative algebraic methods for controllability of discrete-time polynomial systems. *Int. J. Control.*, **89**, 343–351.

Klamka, J. (1993) Controllability of dynamical systems. A survey. *Arch Control Sci*, **2**, 281–307.

Koroljuk, V. S., Borovskich and Yu. V. (1994) *Theory of U-statistics*. Mathematics and its Applications, 273. Kluwer Academic Publishers Group, Dordrecht.

Kusuoka, S. and Tudor, C. A. (2012) Stein’s method for invariant measures of diffusions via Malliavin calculus. *Stoch. Proc. Appl.*, **122**, 1627–1651.

Lee, A. J. (1990) *U-statistics. Theory and practice*. Statistics: Textbooks and Monographs, 110. Marcel Dekker, Inc., New York.

Ley, C., Reinert, G. and Swan, Y. (2017) Stein’s method for comparison of univariate distributions. *Probab. Surv.*, **14**, 1–52.

Ley, C. and Swan, Y. (2013) Stein’s density approach and information inequalities. *Electron. Probab.*, **18**, no. 7, 1–14.

Nourdin, I., Nualart, D. and Peccati, G. (2016) Quantitative stable limit theorems on the Wiener space. *Ann. Probab.*, **44**, 1–41.

Nourdin, I. and Peccati, G. (2009) Stein’s method on Wiener chaos. *Probab. Theory Rel.*, **145**, 75–118.
Nourdin, I. and Peccati, G. (2010) Cumulants on the Wiener space. J. Funct. Anal., 46, 45–58.
Nourdin, I. and Peccati, G. (2012) Normal Approximations Using Malliavin Calculus: from Stein’s Method to Universality. Cambridge Tracts in Mathematics. Cambridge University Press.
Nourdin, I. and Peccati, G. (2015) The optimal fourth moment theorem. P. Am. Math. Soc., 143, 3123–3133.
Nourdin, I., Peccati, G. and Réveillac, A. (2010) Multivariate normal approximation using Stein’s method and Malliavin calculus. Ann. I. H. Poincare-PR, 46, 45–58.
Nourdin, I. and Poly, G. (2012) Convergence in law in the second Wiener/Wigner chaos. Electron. Commun. Probab., 17, no. 36, 1–12.
Nualart, D. and Nualart, E. (2018) Introduction to Malliavin Calculus. In Institute of Mathematical Statistics Textbooks. Cambridge University Press.
Peccati, G. (2014) Quantitative CLTs on a Gaussian space: a survey of recent developments. ESAIM: Proceedings, 44, 61–78.
Peköz, E., Röllin, A. and Ross, N. (2013) Degree asymptotics with rates for preferential attachment random graphs. Ann. Probab., 23, 1188–1218.
Pike, J. (2014) Stein’s method and the Laplace distribution. ALEA Lat. Am. J. Probab. Stat., 11, 571–587.
Ross, N. (2011) Fundamentals of Stein’s method. Probab. Surv., 8, 210–293.
Rugh, W. (1993) Linear system theory. Prentice Hall Information and System Sciences Series.
Schoutens, W. (2001) Orthogonal polynomials in Stein’s method. J. Math. Anal. Appl., 253, 515–531.
Sontag, E. D. (1998) Mathematical Control Theory: Deterministic Finite Dimensional Systems. second ed., Springer, New York.
Stein, C. (1972) A bound for the error in the normal approximation to the the distribution of a sum of dependent random variables. In Proc. Sixth Berkeley Symp. Math. Statist. Probab. vol. 2, University of California Press, Berkeley, 583–602.
Stein, C. (1986) Approximate Computation of Expectations. IMS, Hayward, California.
Stein, C., Diaconis, P., Holmes, S. and Reinert G. (2004) Use of exchangeable pairs in the analysis of simulations. In Stein’s method: expository lectures and applications (P. Diaconis and S. Holmes, eds.), IMS Lecture Notes Monograph Series vol. 46, Beachwood, Ohio, USA: Institute of Mathematical Statistics, 1–26.
Triggiani, R. (1975) Controllability and Observability in Banach Spaces with Bounded Operators. SIAM J. Control, 13, 462–491.
Triggiani, R. (1976) Extensions of Rank Conditions for Controllability and Observability to Banach Space and Unbounded Operators. SIAM J. Control, 14, 313–334.
Xu, L. (2019) Approximation of stable law in Wasserstein-1 distance by Stein’s method. Ann. Appl. Probab., 29, 458–504.