Multidimensional version of Kolmogorov's Theorem.

Let us do $d = 2$. $d > 2$ is not all that different. We need to interpolate a function from the four corners of a square to its interior. Pretending the square to be $[0,1]^2$, the function will be of the form

$$f(x_1, x_2) = a + bx_1 + cx_2 + dx_1x_2$$

It is linear in each variable. The values on the edge of a square are obtained by linear interpolation from the corners. This guarantees that the function defined on each one for the sub-squares separately matches at the common edges and defines a continuous function on the big square. The comparison between the interpolated functions u_n, u_{n+1} at n-th stage and the $n+1$-th stage involves differences at 2^{2n} possible nodes. [This will be 2^{nd} in d-dimensions]. As before

$$\sup_x |u_n(x) - u_{n+1}(x)| \leq \max_{1 \leq i \leq c_d 2^{2n}} |u(x_i) - u(y_i)|$$

with $\sup_i |x_i - y_i| \leq 2^{-n}$. x_i is the mid point of an edge at the n-th stage and y_i is either end point of that edge. At the n-th stage the number of such comparisons can be bounded by $c_d 22n$ where c_d is a simple constant that depnds on the dimension. Now Tchebechev inequality will do the trick provided we have for some $\alpha > 0$,

$$E[|u(x) - u(y)|^\beta] \leq |x - y|^{2+\alpha}$$

Converse Estimate. We now prove the reverse estimate

Theorem.

$$\|f\|_{L_p(P)} \leq C_p \|\Lambda f\|_{1,p}$$

Proof: This is done by duality. We note that

$$< -\mathcal{L}f, g >_{L_2(P)} = \int_X < Df, Dg >_{\mathcal{H}} (x) P(dx)$$

provided f and g are in $\cup_n \mathcal{K}_n$ i.e. are polynomials. Therefore

$$< f|g > = \int_X [< Df(x)\, , Dg(x) >_{\mathcal{H}} + f(x)g(x)] P(dx)$$

defines an inner product on polynomials.

$$E^P[fg] = E^P[(\Lambda^2 f)((I-\mathcal{L})g)] = <\Lambda f\,|\Lambda g>$$
$$= \int_X [<(D\Lambda f)(x)\,,(D\Lambda g)(x)>_{\mathcal{H}} + (\Lambda f)(x)(\Lambda g)(x)]P(dx)$$

If we now take sup over $g \in L_q(P)$, and use the inequalities in the other direction we get the theorem.

Comments on Higher Derivatives.

Suppose $f(\omega)$ is a function on X then the higher derivatives $D^r f$ are symmetric r-linear functional on $\mathcal{H}$ and can be viewed as a symmetric element of the tensor product $\otimes_r \mathcal{H}$ and derives its norm. For instance, for $r=2$, it would be the Hilbert-Schmidt norm, where an element in $\mathcal{H}\otimes\mathcal{H}$ is viewed as a symmetric operator. Now Leibnitz rule applies and

$$D^k(fg) = \sum_{r+s=k} c_{r,s} D^r f \otimes D^s g$$

for some coefficients $c_{r,s}$. In particular if $f, g \in \mathcal{S}$, where

$$\mathcal{S} = \cap_{p,r}\big\{f \in L_p(P), \|D^r f\|_{\otimes_r \mathcal{H}} \in L_p(P)\big\}$$

then $fg \in \mathcal{S}$.

For each k, the norm

$$\|f\|_{k,p} = \sum_{0\le r\le k} \|D^r f\|_{L_p}$$

can be shown to be equivalent to the norm $\|\Lambda^{-k} f\|_{L_p}$ i.e

$$D^k \Gamma^k$$

is bounded from $L_p \to L_p$.

We cannot use directly the estimates for maps into a Hilbert space because our goal is to estimate $D^n\Lambda^n f$ with values in $\otimes_{j=1}^n \mathcal{H}$ in terms of the scalar function f. The idea is to study some intertwining operators and reduce the problem to the boundedness of $D\Lambda$.

The norms for higher derivatives can be defied inductively, in fact for functions with values in some $\mathcal{V}$.

$$\|f\|_{r,p} = \|f\|_{r-1,p} + \|Df\|_{L_p(P,\mathcal{H}^{\otimes(r-1)}\otimes\mathcal{V})}$$

Theorem. There exists a constant $c = c(r, p)$ such that

$$c^{-1}\|f\|_{L_p(P)} \leq \|\Lambda^r f\|_{r,p} \leq c\|f\|_{L_p(P)}$$

Proof: Consider the map $\Gamma : \mathcal{D}_{1,p}(X; \mathcal{V}) \to L_p(P; \mathcal{V} \oplus (\mathcal{H} \otimes \mathcal{V}))$ defined by

$$\Gamma f = (f\,, Df)$$

Basically by induction

$$c_r^{-1}\|f\|_{p,r} \leq \|\Gamma^r f\|_{L_p(P)} \leq c_r\|f\|_{p,r}$$

We construct a map $A_k : L_p(P; \mathcal{V}) \oplus L_p(P; \mathcal{H} \otimes \mathcal{V})) \to L_p(P; \mathcal{V}) \oplus L_p(P; \mathcal{H} \otimes \mathcal{V}))$ so that

$$A_k\Gamma = \Lambda^{-k}\Gamma\Lambda^k$$

We seethat in the chaos decomposition D lowers the degree by 1, so that Γ leaves the first component alone while lowering the degree by one in the second component. Therefore A_k can be taken as I in the first component and as mutiplication by $(\frac{n+1}{n+2})^{\frac{k}{2}}$ on terms of degree n. We see that

$$\begin{aligned}
\Gamma^n\Lambda^n &= \Gamma^{n-1}\Gamma\Lambda^{n-1}\Lambda \\
&= \Gamma^{n-1}\Lambda^{n-1}A_{n-1}\Gamma\Lambda \\
&= (\Gamma\Lambda)A_1(\Gamma\Lambda)A_2\cdots(\Gamma\Lambda)A_{n-1}(\Gamma\Lambda)
\end{aligned}$$

Since at each step A_j and $\Gamma\Lambda$ are bounded operators in every L_p we are done.

Divergence Operator: Given a map $u \in L_p(P\,; \mathcal{H} \otimes \mathcal{V})$ the divergence $v = D^*$ is defined as the map $X \to \mathcal{V}$ defined by

$$\int_X < f(x)\,, v(x) >_{\mathcal{V}} = \int_X < (Df)(x)\,, u(x) >_{\mathcal{H}\otimes\mathcal{V}} P(dx)$$

for all smooth functions f with values in $\mathcal{V}$.

Theorem. For $u \in \mathcal{D}_{1,p}$, $v = D^*u$ exists in $L_p(P; \mathcal{V})$. More precisely there is a constant c_p such that

$$\|D^*u\|_{L_p(P;\mathcal{V})} \leq c_p\|u\|_{1,p}$$

Proof:

Commutation relations.

$$DP_t = e^{-t} P_t \, D$$

$$D\Lambda^{-1} = M\Lambda^{-1}D$$

where

$$M\,f = \sqrt{\frac{n+1}{n+2}}\,f$$

on $\mathcal{K}_n$.

Riesz Transform. If we define $R = D\,\Lambda : L_p(P) \to L_p(P;\mathcal{H})$, then

$$\|R\,f\|_{L_p(P;\mathcal{H})} \leq c_p \|f\|_{L_p(P)}$$

We also have

$$D = R\Lambda^{-1} \qquad \text{and} \qquad D = \Lambda^{-1}MR$$

Finally,

$$\begin{aligned} \int_X < (Df)(x)\,, u(x) >_{\mathcal{H}} P(dx) &= \int_X < (\Lambda^{-1}MR)\,f(x)\,, u(x) >_{\mathcal{H}} P(dx) \\ &= \int_X f(x)\,(R^* M\Lambda^{-1})u(x)\,P(dx) \end{aligned}$$

Therefore

$$D^* u = R^* M\Lambda^{-1} u$$

and satisfies the bound

$$\|D^* u\|_{L_p(P)} \leq c_p \|u\|_{1,p}\,.$$

We can think of a map $A : X \to H$ as a vector field and its divergence

$$\delta A = D^* A$$

satisfies

$$\|\delta A\|_{L_p(P)} \leq c_p \|A\|_{1,p}\,.$$

Malliavin Covariance Matrix. Suppose $g(x) : X \to R^d$ is a map with $\|g\|_{1,p} < \infty$. Then

$$(Dg)(x) \in \mathcal{H} \otimes R^d \quad \text{a.e.} \quad P$$

or representing $g = \{g_i\}$ we have for each $i = 1, \cdot, d$

$$(Dg_i)(x) \in \mathcal{H} \quad \text{a.e.} \quad P$$

The Malliavin Covariance Matrix is the symmetric positive semi-definite matrix

$$\sigma(x) = \sigma_{i,j}(x) =< (Dg_i)(x)\,,(Dg_j)(x) >_{\mathcal{H}}$$

exists and is in $L_{\frac{p}{2}}(P)$ provided $p \geq 2$. The map g is called *non degenerate* if

$$[\det \sigma(x)]^{-1} \in L_p(P)$$

for every $1 \leq p < \infty$. It is called *weakly non degenerate* if

$$\det \sigma(x) > 0 \qquad \text{a.e.} \quad P$$

The map $g : X \to R^d$ defines a gradient $(Dg)(x)$ which is a linear map $g'(x)$ from the tangent space $\mathcal{H}$ of X at x to the tangent space R^d of R^d at $g(x)$. Then $\sigma(x) = g'(x)g'^{\,*}(x)$. Given a map g and a vector field $z = z(y)$ on R^d i.e. a map $R^d \to R^d$, we can look for a vector field $\tilde{Z}$ on X a *lift* of z such that

$$g'(x)\tilde{Z}(x) = z(g(x))$$

In the nondegenerate case this is possible, at least for almost all x. A canonical choice which, for each x, minimizes $\|\tilde{Z}(x)\|_{\mathcal{H}}$ is given by

$$Z(x) = g'^{\,*}(x)[\sigma(x)]^{-1}\, z(g(x))$$

In particular we can lift $\frac{\partial}{\partial y_k}$ to

$$Z_k(x) = \sum_j \gamma_{k,j}(x) g'_j(x)$$

where $\gamma(x)$ is the inverse of $\sigma(x)$.

Smoothness of distributions. Let $g(x)$ be a map into R^d that is nondegenerate and smooth in the sense that $\|g\|_{r,p} < \infty$ for all r and p. If $f(y)$ is a smooth function on R^d, and $\rho(dy)$ is the distribution $\rho = Pg^{-1} = g_*P$ we have

$$\begin{aligned}\int_{R^d} \frac{\partial f}{\partial y_k}(y)\rho(dy) &= \int_X < D\tilde{f}\,, Z_k > (x) dP \\ &= \int_X \tilde{f}(x)\,(\delta Z_k)(x)\, dP \\ &= \int_{R^d} f(y) v_k(y) \rho(dy)\end{aligned}$$

where $\tilde{f} = f(g(x))$ is the lifted function f and $v_k(y) = E^P[\delta Z_k | g(\cdot)]$ is the conditional expectation. If we can get estimates on $\|\delta Z_k\|_{L_p(P)}$ that will be fine because

$$\int_{R^d} \Big|\frac{\partial r}{\partial y_k}(y)\Big|^p r(y)\, dy = \int_{R^d} |v_k(y)|^p \rho(dy) \leq \int_X |\delta Z_k(x)|^p P(dx)$$

where $\rho(dy) = r(y)dy$.

Calculation of δZ_k. From the definition

$$Z_k = \sum_j \gamma_{k,j}(x) g'_j(x)$$

we can compute

$$(\delta Z_k)(x) = \sum_j < [D\,\gamma_{k,j}](x)\,,(Dg_j)(x) > + \sum_j \gamma_{k,j}(x)(\delta D g_j)(x)$$

and using the relations

$$\delta D = -\mathcal{L} \quad \text{and} \quad D\gamma = D\sigma^{-1} = \sigma^{-1}(Da)\sigma^{-1} = \gamma(Da)\gamma$$

we can write

$$\begin{aligned}
(\delta Z_k)(x) &= -\sum_j \gamma_{k,j}(x)(\mathcal{L}g_j)(x) + \sum_{s,j,i} \gamma_{k,s}(x)\gamma_{j,i}(x) < (D\sigma_{s,j})(x)\,,(Dg_i)(x) > \\
&= -\sum_j \gamma_{k,j}(x)(\mathcal{L}g_j)(x) \\
&\qquad + \sum_{s,j,i} \gamma_{k,s}(x)\gamma_{j,i}(x) < (D < (Dg_s)(\cdot)\,,(Dg_j)(\cdot) >)(x)\,,(Dg_i)(x) > \\
&= -\sum_j \gamma_{k,j}(x)(\mathcal{L}g_j)(x) + \sum_{s,j,i} \gamma_{k,s}(x)\gamma_{j,i}(x)\Big[(D^2 g_s)(x)[(Dg_j)(x)\,,(Dg_i)(x)]\Big] \\
&\quad + \sum_{s,j,i} \gamma_{k,s}(x)\gamma_{j,i}(x)\Big[(D^2 g_j)(x)[(Dg_s)(x)\,,(Dg_i)(x)]\Big] \\
&= -\sum_j \gamma_{k,j}(x)(\mathcal{L}g_j)(x) + \sum_{s,i} \gamma_{k,s}(x)\big[(D^2 g_s)(Z_i\,,Dg_i)\big](x) \\
&\quad + \sum_j \big[(D^2 g_j)(Z_k\,,Z_j)\big](x)
\end{aligned}$$

Since terms of the form

$$\big[(D^2 g)(u\,,v)\big](x)$$

can be estimated by

$$\|(D^2 g)(x)\|_{\mathcal{H}\otimes\mathcal{H}} \times \|u(x)\|_{\mathcal{H}} \times \|v(x)\|_{\mathcal{H}}$$

we acn control $\|\delta Z_k\|_{L_p(P)}$ by $\|\gamma\|_{L_{p'}}(P)$ and $\|g\|_{2,p'}$ with large enough p'.