Lecture 11

*Lecture date: October 4, 2004* *Scribe: Elchanan Mossel*

# 1 Hyper-contraction of noise operators

In this section we begin the analysis of *noise-correlation.* The main interest here is understanding the correlation between $f(x_1, \ldots, x_n)$ and $f(y_1, \ldots, y_n)$ where $(x_1, \ldots, x_n)$ are chosen from a product distribution and $(y_1, \ldots, y_n)$ is obtained from $(x_1, \ldots, x_n)$ be applying some noise to each coordinate independently. The main difference in our study here compared to the study of influences will be our interest in re-randonmizing many coordinates simultanously, instead of studying the perturbation caused by a single parameter. Interestingly, our first application of this theory of noise-correlation will be to the study of influences.

We begin with a general defintion of tensor product of operators – this corresponds to applying noise independently to each coordinate. Then we will study a strong propery of these operators, named hyper-contraction – this will be use frequently later.

## 1.1 Noise operators

**Definition 1** *A operator $T : L^2(\mu) \to L^2(\mu)$ is called positivity improving if $Tf \geq 0$ for all $f \geq 0$. We will call $T$ a* noise operator *if it is positivity improving, $\|Tf\|_2 \leq \|f\|_2$ for all $f$, $T1 = 1$ and $\langle Tf, g\rangle = \langle f, Tg\rangle$ for all $f, g \in L^2(\mu)$.*

**Example 2** *Let $(\Omega, \mu)$ be a finite probability space and let $M$ a Markov chain that is reversible with respect to $\mu$. $M$ corresponds to a non-negative $|M| \times |M|$ matrix that satisfies:*

$$\sum_y M(x,y) = 1,$$

*for all $x \in \Omega$ and*

$$\mu(x)M(x,y) = \mu(y)M(y,x)$$

*for all $x$ and $y$. Let $T_M$ be defined as follows*

$$(T_M f)(x) = \sum_y M(x,y)f(y).$$

*Then $T_M$ a is noise operator:*

$$\begin{aligned}
\|T_M f\|_2^2 &= \sum_x \mu(x)(T_M f)^2(x) = \sum_{x,y} \mu(x)(M(x,y)f(y))^2 \\
&\le \sum_{x,y} \mu(x)M(x,y)(f(y))^2 = \sum_y \mu(y)f^2(y) = \|f\|_2^2,
\end{aligned}$$

$$\langle T_M f, g\rangle = \sum_x \mu(x)T_M f(x)g(x) = \sum_{x,y} \mu(x)M(x,y)f(y)g(x) = \sum_{x,y} \mu(y)M(y,x)f(y)g(x) = \langle f, T_M g\rangle.$$

**Example 3** *Consider the space $L^2(\gamma_n)$ where $\gamma_n$ is the $n$-dimensional Gaussian measure. Let $0 \le \rho \le 1$. The* Orenstein-Uhelenbeck *operator is defined by:*

$$T_\rho f(x) = \mathbf{E}_{y\sim\gamma_n}[f(\rho x + \sqrt{1-\rho^2}y)].$$

*In order to check that this is a noise operator note that*

$$\begin{aligned}
\mathbf{E}_{x\sim\gamma_n}[(T_\rho f)^2(x)] &= \mathbf{E}_{x\sim\gamma_n}\left[\mathbf{E}^2_{y\sim\gamma_n}[f(\rho x + \sqrt{1-\rho^2}y)|x]\right] \\
&\le \mathbf{E}_{x\sim\gamma_n, y\sim\gamma_n}[f^2(\rho x + \sqrt{1-\rho^2}y)] = \mathbf{E}_{x\sim\gamma_n}[f^2(x)],
\end{aligned}$$

*where the last equality follows from the fact that if $N_1, N_2$ are two independent standard Gaussian vectors, then so is $\rho N_1 + \sqrt{1-\rho^2}N_2$.*

*We also have that*

$$\langle T_\rho f, g\rangle = \mathbf{E}[f(X)f(Y)],$$

*where $(X,Y)$ is a normal $2n$-dimensional vector where $\mathbf{Cov}[X_i, X_j] = \mathbf{Cov}[Y_i, Y_j] = \delta_{i,j}$ and $\mathbf{Cov}[X_i, Y_j] = \rho\delta_{i,j}$. Since this expression is symmetric in $X$ and $Y$ it follows that*

$$\langle T_\rho f, g\rangle = \langle f, T_\rho g\rangle.$$

## 1.2 Tensor products of noise operators

**Definition 4** *Let $T_i : L^2(\mu_i) \to L^2(\mu_i)$ be a bounded linear operator. Let $U^i$ be a basis of $L^2(\mu_i)$. We define $T = \otimes_{i=1}^n T_i$ to be the linear operator satisfying*

$$T(\otimes_{i=1}^n u_i) = \otimes_{i=1}^n (T_i u_i),$$

*for every basis element $\otimes_{i=1}^n u_i$.*

This definition roughly says that $T$ acts on coordinates $i$ by $T_i$. One needs to check that this definition does not depend on the choice of basis.

**Lemma 5** *The operator $T$ does not depend on the choice of basis.*

**Proof:** We need to prove that for any two bases $\otimes_{i=1}^n U^i$ and $\otimes_{i=1}^n V^i$ we get the same operator. Clearly it suffices to show that assuming $U^i = V^i$ except at a single coordinate $i$ that we may assume WLOG is 1. In other words, it suffices to show we obtain the same operator for $U^1 \otimes \ldots \otimes U^n$ and for $V^1 \otimes U^2 \otimes U^n$. This follows immediately from the linearity of $T_1$. $\square$

**Lemma 6** *Let $T_i : L^2(\mu_i) \to L^2(\mu_i)$ be a bounded linear operators. Let $T_i^* : L^2(\prod_{i=1}^n \mu_i) \to L^2(\prod_{i=1}^n \mu_i)$ be defined by*

$$(T_i^* f(\cdot, \ldots, \cdot))(x_1, \ldots, x_n) = (T_i f(x_1, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_n))(x_i).$$

*Then $\prod_{i=1}^n T_i^* = \otimes_{i=1}^n T_i$ and the operators $T_i^*$ commute.*

**Proof:** It suffices to check that $\prod_{i=1}^n T_i^* = \prod_{i=1}^n T_i$ for basis elements. $\square$

**Lemma 7** *If $T_1, \ldots, T_n$ are noise operators then so is $\otimes_{i=1}^n T_i$.*

**Proof:** It is easy to see that each of the $T_i^*$ is a noise operator. $\square$

**Lemma 8** *Suppose $T^i$ is a Markov operator on $L^2(\mu_i)$ that is defined by a reversible Markov chain $M^i$. Then the operator $\otimes_{i=1}^n T^i$ is the operators defined by the Markov chain $M$ where,*

$$M(x, y) = \prod_{i=1}^n M^i(x_i, y_i).$$

**Proof:** It suffices to show that the two operators acts the same on tensors. Let $u = \otimes_{i=1}^n u_i$ be such a tensor then

$$(T_M u)(x) = \sum_y M(x, y) u(y) = \sum_y \prod_{i=1}^n M^i(x_i, y_i) u_i(y_i) = \prod_{i=1}^n \left( \sum_{y_i} M^i(x_i, y_i) u_i(y_i) \right) = \prod_{i=1}^n T_{M^i}(u_i),$$

as needed. $\square$

**Example 9** *The most important noise operator we will study is the* Bonami-Beckner *operator. This operators is specified by a single paramter $0 \le \rho \le 1$. The operator $T_\rho$ is defined on $L^2(\prod_{i=1}^n \mu_i)$ by $T_\rho = \otimes_{i=1}^n T_\rho^i$, where $T_\rho^i(f) = \rho f + (1-\rho)\mathbf{E}[f]$. Note that the operator $T^i$ may be defined via the Markov chain $M^i$ where $M^i(x, y) = \rho \delta_x + (1-\rho)\mu(y)$. Therefore the operator $T_\rho$ corresponds to $M(x, y)$ where $y_i = x_i$ with probability $\rho$ and is chosen independently from the measure $\mu$ indpenedently for all $i$.*

Noise operators are contractions by definition. They satisfy $\|Tf\|_2 \leq \|f\|_2$ . More importantly, many of these operators are hyper-contractive.

**Definition 10** *Let $1 \leq p \leq q$ then we say that the operator $T$ is $(p,q)$-*hypercontractive *satisfies $\|Tf\|_q \leq \|f\|_p$ for every $f$ with $\|f\|_p < \infty$.*