FOR NO NOISE

$$\hat{r} = \frac{r^2 + r}{r+1} = r$$

$$\hat{r} = h(\underline{w}) = \frac{r^2 + w[2] + r + w[1]}{r + w[1] + 1 + w[0]}$$

$$E(\hat{r}) = h(\underline{0}) = r \qquad \text{UNBIASED}$$

$$\left.\frac{\partial h}{\partial w[0]}\right|_{\underline{w}=\underline{0}} = - \left.\frac{r^2 + w[2] + r + w[1]}{(r + w[1] + 1 + w[0])^2}\right|_{\underline{w}=\underline{0}}$$

$$= \frac{-r}{r+1}$$

SIMILARLY, $\left.\frac{\partial h}{\partial w[1]}\right|_{\underline{w}=\underline{0}} = - \frac{r-1}{r+1}$

$$\left.\frac{\partial h}{\partial w[2]}\right|_{\underline{w}=\underline{0}} = \frac{1}{r+1}$$

BUT $\underline{C} = \sigma^2 \underline{I}$

$$\Rightarrow \quad \mathrm{VAR}(\hat{r}) = \sigma^2 \sum_{n=0}^{2} \left( \left.\frac{\partial h}{\partial w[n]}\right|_{\underline{w}=\underline{0}} \right)^2$$

$$= 2\sigma^2 \, \frac{r^2 - r + 1}{(r+1)^2}$$

NOTE CRLB IS $\frac{\sigma^2}{1+4\sigma^2}$.

---

FINALLY, NOTE THAT IF $\underline{I}$ IS GAUSSIAN AS $N \to \infty$ OR $\underline{W}$ IS GAUSSIAN $\Rightarrow \hat{\theta}$ WILL BE ASYMPTOTICALLY GAUSSIAN. WHY?

## BAYESIAN ESTIMATION

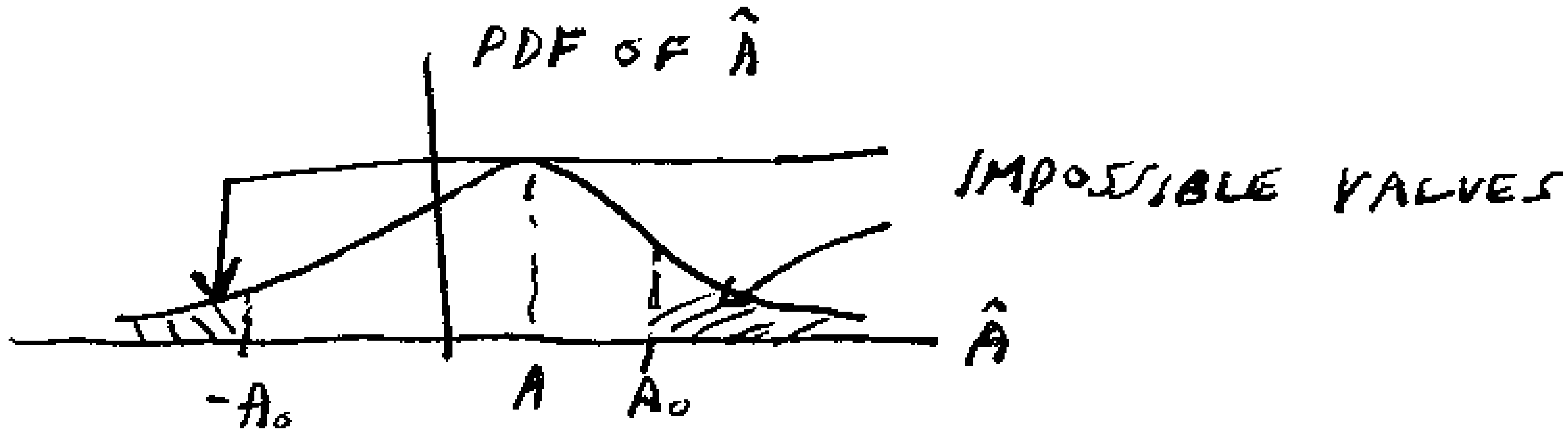
NOW ASSUME WE HAVE SOME PRIOR KNOWLEDGE ABOUT $\theta$. TO INCORPORATE IT ASSUME $\theta$ IS RANDOM VARIABLE WITH GIVEN PDF AND ESTIMATE $\theta$ AS REALIZATION OF RANDOM VARIABLE.

### PRIOR KNOWLEDGE

EXAMPLE - DC LEVEL IN WGN

BUT $-A_0 \le A \le A_0$ INSTEAD OF $-\infty < A < \infty$

$\hat{A} = \bar{x}$ STILL MVU BUT WE CAN DO BETTER



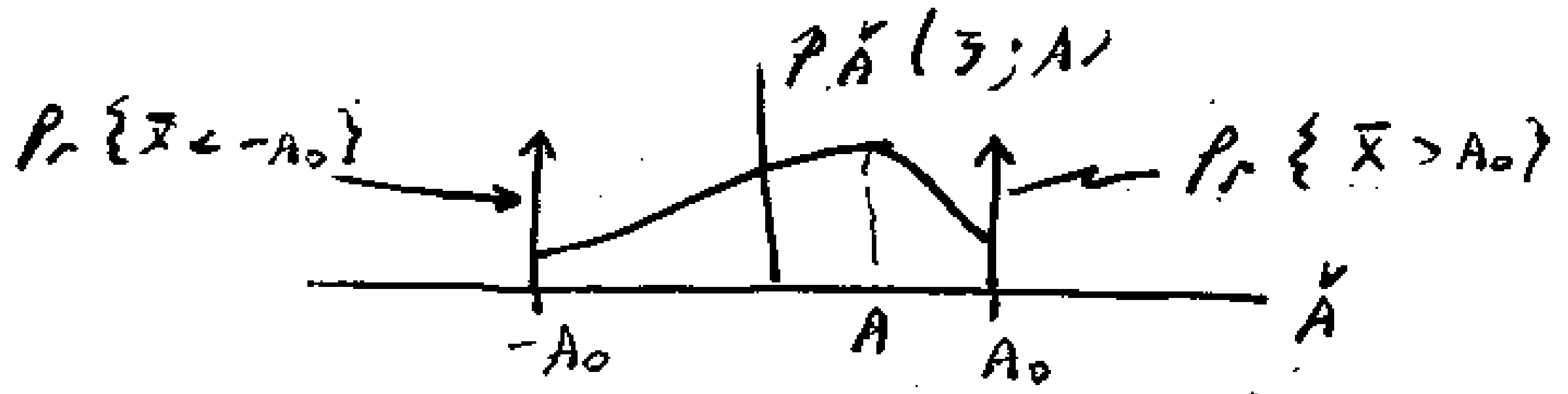
$|\hat{A}| > A_0$ LEADS TO IMPOSSIBLE ESTIMATES FOR $A$

CONSIDER TRUNCATED SAMPLE MEAN (BIASED)

$$\check{A} = \begin{cases} -A_0 & \bar{x} < -A_0 \\ \bar{x} & -A_0 \le \bar{x} \le A_0 \\ A_0 & \bar{x} > A_0 \end{cases}$$

CONSISTENT WITH KNOWN CONSTRAINTS

COMPARE MSES OF $\hat{A}$ AND $\check{A}$




$$p_{\check{A}}(\xi;A) = Pr\{\bar{x} < -A_0\}\,\delta(\xi + A_0) + p_{\hat{A}}(\xi;A)\left[u(\xi + A_0) - u(\xi - A_0)\right] + Pr\{\bar{x} > A_0\}\,\delta(\xi - A_0)$$

$$mse(\hat{A}) = \int_{-\infty}^{\infty} (\xi - A)^2 p_{\hat{A}}(\xi;A)\,d\xi$$

$$= \int_{-\infty}^{-A_0} (\xi - A)^2 p_{\hat{A}}(\xi;A)\,d\xi + \int_{-A_0}^{A_0} (\xi - A)^2 p_{\hat{A}}(\xi;A)^2\,d\xi + \int_{A_0}^{\infty} (\xi - A)^2 p_{\hat{A}}(\xi;A)\,d\xi$$

$$> \int_{-\infty}^{-A_0} (-A_0 - A)^2 p_{\hat{A}}(\xi;A)\,d\xi + \int_{-A_0}^{A_0} (\xi - A)^2 p_{\hat{A}}(\xi;A)\,d\xi + \int_{A_0}^{\infty} (A_0 - A)^2 p_{\hat{A}}(\xi;A)\,d\xi = mse(\check{A})$$

FOR $-A_0 \leq A \leq A_0$.

$\Rightarrow$ $\check{A}$ BETTER THAN $\hat{A} = \bar{x}$ IN TERMS OF MSE

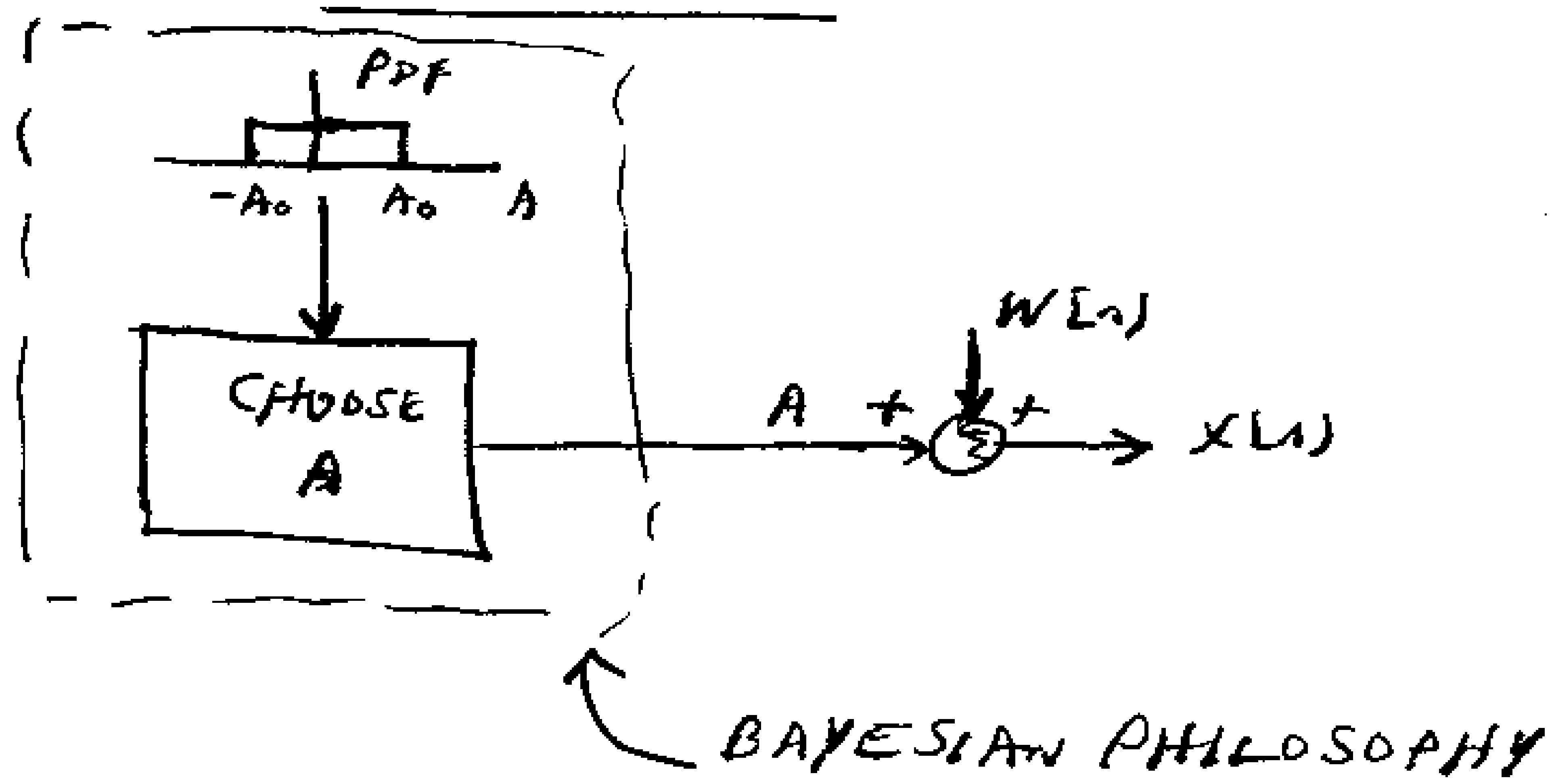
SINCE WE HAVE FOUND A BETTER ESTIMATOR CAN WE FIND AN OPTIMAL ONE. (MINIMUM MSE).

APPROACH: 1) ASSUME THAT A HAS BEEN CHOSEN FROM $[-A_0, A_0]$ INTERVAL

2) CONSIDER CHOOSING PROCESS AS RANDOM EVENT

3) ASSIGN PDF TO A, $U[-A_0, A_0]$ SINCE A IS JUST AS "LIKELY" TO BE ANYWHERE IN INTERVAL

BAYESIAN APPROACH - A CONSIDERED TO BE A RANDOM VARIABLE WITH PRIOR PDF, WE ATTEMPT TO ESTIMATE THE REALIZATION OF A



OPTIMALITY CRITERION IS BAYESIAN MSE

$$Bmse(\hat{A}) = E[(\hat{A}-A)^2]$$

↑ WITH RESPECT TO $p(\underline{x}, A)$

$p(\underline{x}, A)$ = <u>JOINT</u> PDF

CLASSICAL: $MSE(\hat{A}) = \int (\hat{A}-A)^2 p(\underline{x}; A)\, d\underline{x}$

BAYESIAN: $BMSE(\hat{A}) = \int (\hat{A}-A)^2 p(\underline{x}, A)\, d\underline{x}\, dA$

MSE DEPENDS ON A, BMSE DOES NOT

BASIC DATA GENERATING MECHANISM DIFFERENT

COMPUTER EXPERIMENT:

1) CLASSICAL - GENERATE M REALIZATIONS OF $w[n]$, ADD TO EACH ONE <u>GIVEN</u> A

2) BAYESIAN - GENERATE $w[n]$, THEN GENERATE A FROM $U[-A_0, A_0]$, ASSUMING A IS INDEPENDENT OF $w[n]$, REPEAT PROCEDURE M TIMES

MSE WILL DEPEND ON A, BMSE WILL NOT

ADOPT BMSE AS OPTIMALITY CRITERION - WILL NOT LEAD TO $\hat{A}$ WHICH IS A FUNCTION OF A.

TO MINIMIZE

$$\mathrm{BMSE}(\hat{A}) = \int (\hat{A}-A)^2 \, p(\underline{x}, A) \, d\underline{x} \, dA$$

$$= \int \left[ \int (\hat{A}-A)^2 \, p(A|\underline{x}) \, dA \right] p(\underline{x}) \, d\underline{x}$$

MINIMIZE FOR GIVEN $\underline{x}$ (SINCE $p(\underline{x}) \geq 0$)

$$\frac{\partial}{\partial \hat{A}} \int (\hat{A}-A)^2 \, p(A|\underline{x}) \, dA =$$

$$\int 2(\hat{A}-A) \, p(A|\underline{x}) \, dA = 0$$

$$\Rightarrow \quad \hat{A} = \int A \, p(A|\underline{x}) \, dA \qquad \left( \int p(A|\underline{x}) \, dA = 1 \right)$$

$$\therefore \quad \hat{A} = E(A|\underline{x})$$

MEAN OF <u>POSTERIOR</u> <u>PDF</u> $p(A|\underline{x})$ MINIMIZES BMSE.

NOTE: $p(A)$ IS <u>PRIOR</u> <u>PDF</u> (BEFORE DATA OBSERVED)

CONTINUING EXAMPLE

$$p(A|\underline{x}) = \frac{p(\underline{x}|A)\,p(A)}{p(\underline{x})}$$

BAYES' RULE (WHY PROCEDURE TERMED BAYESIAN APPROACH)

$$= \frac{p(\underline{x}|A)\,p(A)}{\underbrace{\int p(\underline{x}|A)\,p(A)\,dA}_{\text{NORMALIZING FACTOR}}}$$

$$\Rightarrow \int p(A|\underline{x})\,dA = 1$$

BUT $w[n]$ IS ASSUMED INDEPENDENT OF $A$

$$\Rightarrow p_x(x[n]|A) = p_w(x[n]-A|A)$$
$$= p_w(x[n]-A)$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{1}{2\sigma^2}(x[n]-A)^2}$$

$$p(\underline{x}|A) = \frac{1}{(2\pi\sigma^2)^{N/2}}\, e^{-\frac{1}{2\sigma^2}\sum_n (x[n]-A)^2}$$

IDENTICAL FORM AS $p(\underline{x};A)$ IN CLASSICAL CASE - BUT THIS IS A CONDITIONAL PDF (WOULD NOT BE THE SAME IF $A$ WERE NOT INDEPENDENT OF $w[n]$)

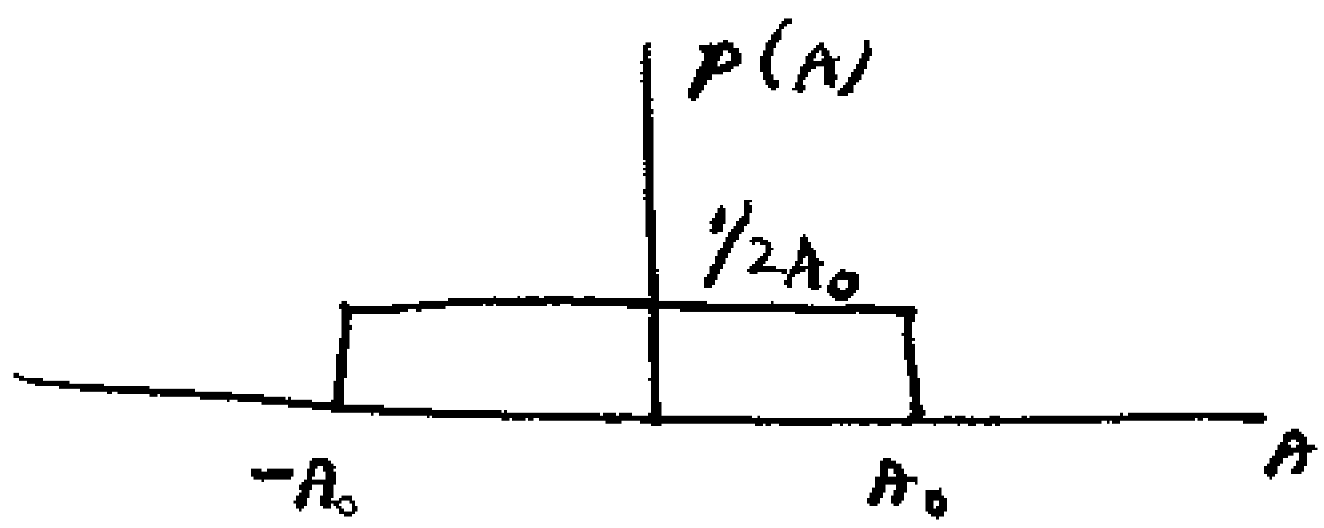
$$p(A|\underline{x}) = \begin{cases} \dfrac{\dfrac{1}{2A_0(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum_n (x[n]-A)^2}}{\int_{-A_0}^{A_0} \quad '' \quad dA} & |A| \le A_0 \\ 0 & |A| > A_0 \end{cases}$$

BUT $\sum (x[n]-A)^2 = \sum x^2[n] - 2NA\bar{x} + NA^2$

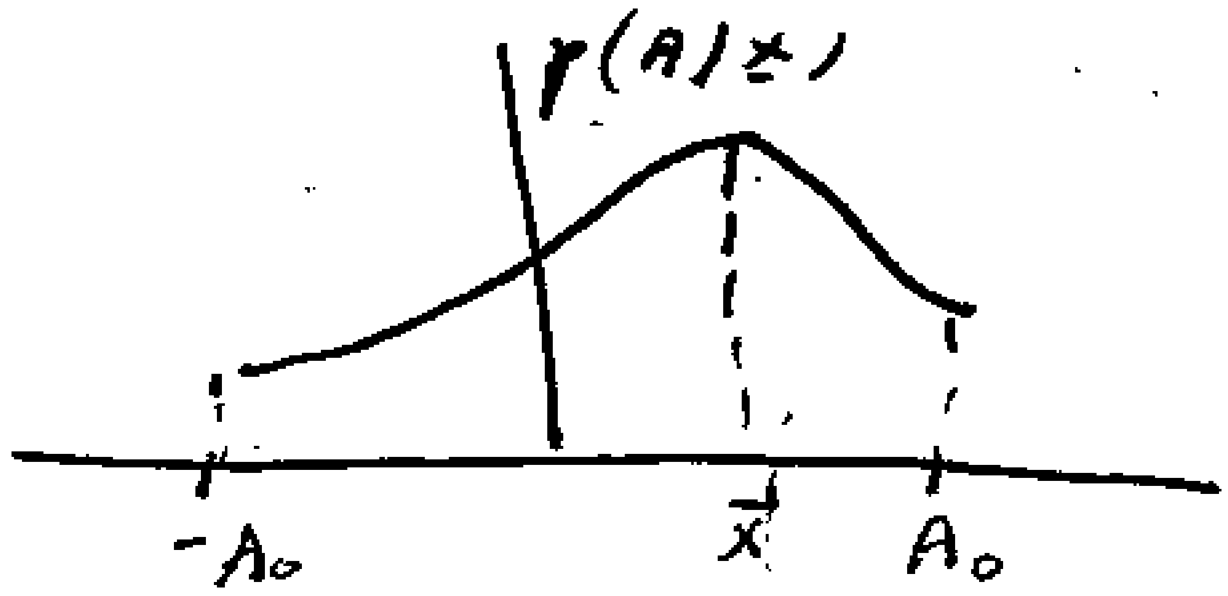
$$p(A|\underline{x}) = \begin{cases} \dfrac{\dfrac{1}{2A_0(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum x^2[n]} e^{-\frac{1}{2\sigma^2}\underbrace{(NA^2-2NA\bar{x})}_{N(A-\bar{x})^2-N\bar{x}^2}}}{\int_{-A_0}^{A_0} \quad '' \quad dA} & |A| \le A_0 \\ 0 & |A| > A_0 \end{cases}$$

$$= \begin{cases} \dfrac{c}{\sqrt{2\pi\sigma^2/N}} e^{-\frac{1}{2\sigma^2/N}(A-\bar{x})^2} & |A| \le A_0 \\ 0 & |A| > A_0 \end{cases}$$

$c$ = NORMALIZING CONSTANT (IF $A_0 \to \infty$, $c = 1$)




PRIOR PDF

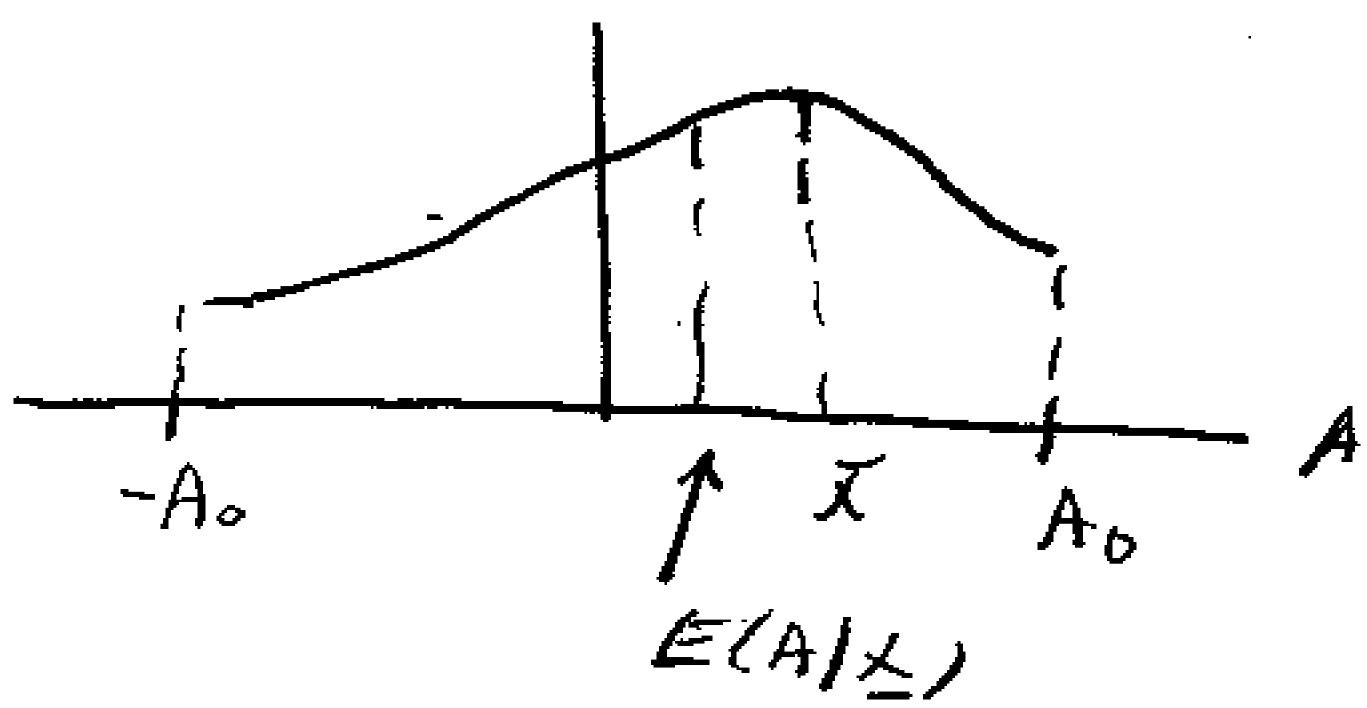



POSTERIOR PDF

NOW $\hat{A} = E(A|\underline{x})$

$$= \frac{\int_{-A_0}^{A_0} A \frac{1}{\sqrt{2\pi\sigma^2/N}} e^{-\frac{1}{2\sigma^2/N}(A-\bar{x})^2} dA}{\underbrace{\int_{-A_0}^{A_0} \frac{1}{\sqrt{2\pi\sigma^2/N}} e^{-\frac{1}{2\sigma^2/N}(A-\bar{x})^2} dA}_{=\frac{1}{c}}}$$

CAN'T BE EVALUATED IN CLOSED FORM.




$\hat{A}$ IS "BIASED" TOWARD ZERO. UNLESS $A_0 \gg \sqrt{\sigma^2/N}$ (NO TRUNCATION OR NO PRIOR KNOWLEDGE)

NOTE THAT BEFORE DATA OBSERVED ALL WE KNOW ABOUT A IS IN $p(A)$.

CAN SHOW THAT

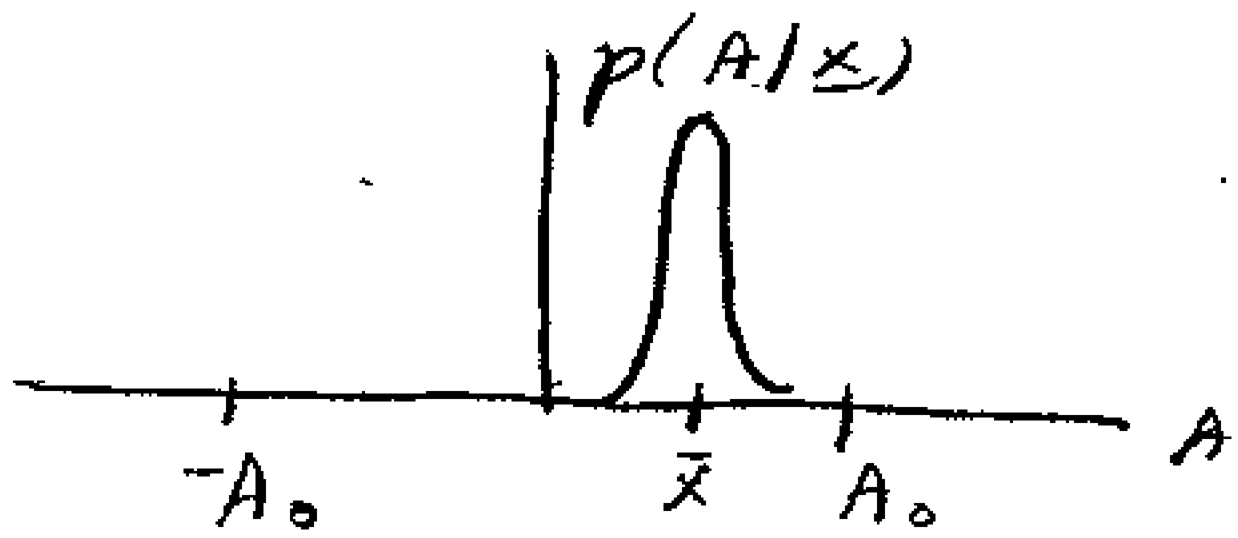
$$E((\hat{A}-A)^2) = \int (\hat{A}-A)^2 p(A)\,dA$$

WHERE $\hat{A}$ NOW A CONSTANT IS MINIMIZED FOR $\hat{A} = E(A)$.

FOR OUR EXAMPLE $\hat{A} = E(A) = 0$

$\Rightarrow$ BAYESIAN ESTIMATOR $E(A|\underline{x})$ IS COMPROMISE BETWEEN PRIOR KNOWLEDGE ESTIMATOR ( $\hat{A} = E(A)$ ) AND "DATA" KNOWLEDGE ESTIMATOR ( $\hat{A} = \bar{x}$ ).

AS $N \to \infty$




$\Rightarrow \hat{A} = E(A|\underline{x}) = \bar{x}$

DATA KNOWLEDGE "SWAMPS OUT" PRIOR KNOWLEDGE.

IN GENERAL,

$$\hat{\theta} = E(\theta|\underline{x})$$

$$= \int \theta\, p(\theta|\underline{x})\,d\theta$$

# CHOOSING PRIOR PDF

SHOULD BE BASED ON PHYSICAL CONSTRAINTS OF PROBLEM IF POSSIBLE.

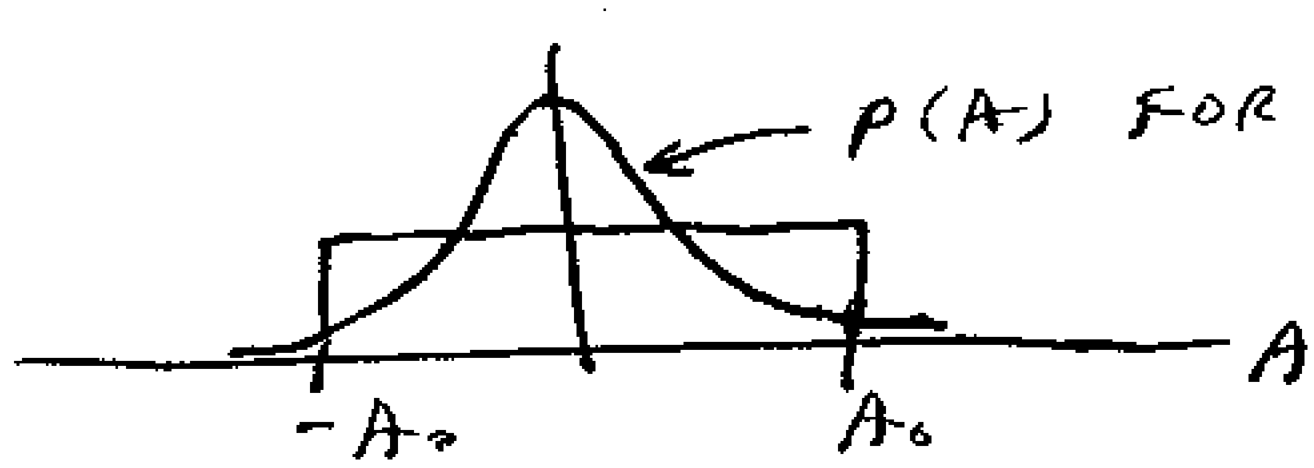
ALSO, CHOOSE PRIOR TO ALLOW EASY INTEGRATION

1) TO FIND $p(\underline{x}) = \int p(\underline{x} \mid \theta) p(\theta) d\theta$

2) TO FIND $E(\theta \mid \underline{x}) = \int \theta \, p(\theta \mid \underline{x}) d\theta$

EXAMPLE - DC LEVEL IN WGN - GAUSSIAN PRIOR PDF

USE $p(A) = \frac{1}{\sqrt{2\pi\sigma_A^2}} e^{-\frac{1}{2\sigma_A^2}(A-\mu_A)^2}$




STRONG BELIEF THAT A WILL BE NEAR $\mu_A$ OR IN INTERVAL $(\mu_A - 3\sigma_A, \mu_A + 3\sigma_A)$

$$p(\underline{x} \mid A) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum (x[n]-A)^2}$$

$$= \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum x^2[n]} e^{-\frac{1}{2\sigma^2}(NA^2 - 2NA\bar{x})}$$

$$p(A|\underline{x}) = \frac{p(\underline{x}|A)\,p(A)}{\int p(\underline{x}|A)\,p(A)\,dA}$$

$$= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}\sqrt{2\pi\sigma_A^2}}\, e^{-\frac{1}{2\sigma^2}\sum x^2(n)}\, e^{-\frac{1}{2\sigma^2}(NA^2 - 2NA\bar{x})} \cdot e^{-\frac{1}{2\sigma_A^2}(A-\mu_A)^2}}{\int_{-\infty}^{\infty} \text{"} \, dA}$$

$$= \frac{e^{-\frac{1}{2\sigma^2}(NA^2 - 2NA\bar{x})}\, e^{-\frac{1}{2\sigma_A^2}(A-\mu_A)^2}}{\int_{-\infty}^{\infty} \text{"} \, dA}$$

$$= \frac{e^{-\frac{1}{2}Q(A)}}{\int_{-\infty}^{\infty} e^{-\frac{1}{2}Q(A)}\,dA}$$

NOTE THAT DENOMINATOR DOES NOT DEPEND ON A., Q(A) IS QUADRATIC IN A $\Rightarrow$ $p(A|\underline{x})$ MUST BE GAUSSIAN.

$$Q(A) = \frac{N}{\sigma^2}A^2 - \frac{2N}{\sigma^2}A\bar{x} + \frac{A^2}{\sigma_A^2} - \frac{2\mu_A A}{\sigma_A^2} + \frac{\mu_A^2}{\sigma_A^2}$$

$$= \underbrace{\left(\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}\right)}_{\frac{1}{\sigma_{A|x}^2}} A^2 - 2\underbrace{\left(\frac{N}{\sigma^2}\bar{x} + \frac{\mu_A}{\sigma_A^2}\right)}_{\frac{\mu_{A|x}}{\sigma_{A|x}^2}} A + \frac{\mu_A^2}{\sigma_A^2}$$

COMPLETING SQUARE

$$Q(A) = \frac{1}{\sigma_{A|x}^2}\left(A^2 - 2\mu_{A|x}A + \mu_{A|x}^2\right) + \frac{\mu_A^2}{\sigma_A^2} - \frac{\mu_{A|x}^2}{\sigma_{A|x}^2}$$

$$= \frac{1}{\sigma_{A|x}^2}\left(A - \mu_{A|x}\right)^2 + \frac{\mu_A^2}{\sigma_A^2} - \frac{\mu_{A|x}^2}{\sigma_{A|x}^2}$$

AFTER THE NORMALIZATION

$$p(A|\underline{x}) = \frac{1}{\sqrt{2\pi\sigma_{A|x}^2}}\, e^{-\frac{1}{2\sigma_{A|x}^2}(A-\mu_{A|x})^2}$$

WHERE $$\sigma_{A|x}^2 = \frac{1}{\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}}$$

$$\mu_{A|x} = \frac{\frac{N}{\sigma^2}\bar{x} + \frac{\mu_A}{\sigma_A^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}}$$

POSTERIOR PDF ALSO GAUSSIAN BUT WITH DIFFERENT MEAN AND VARIANCE.

MMSE ESTIMATOR IS

$$\hat{A} = E(A|\underline{x}) = \mu_{A|x}$$

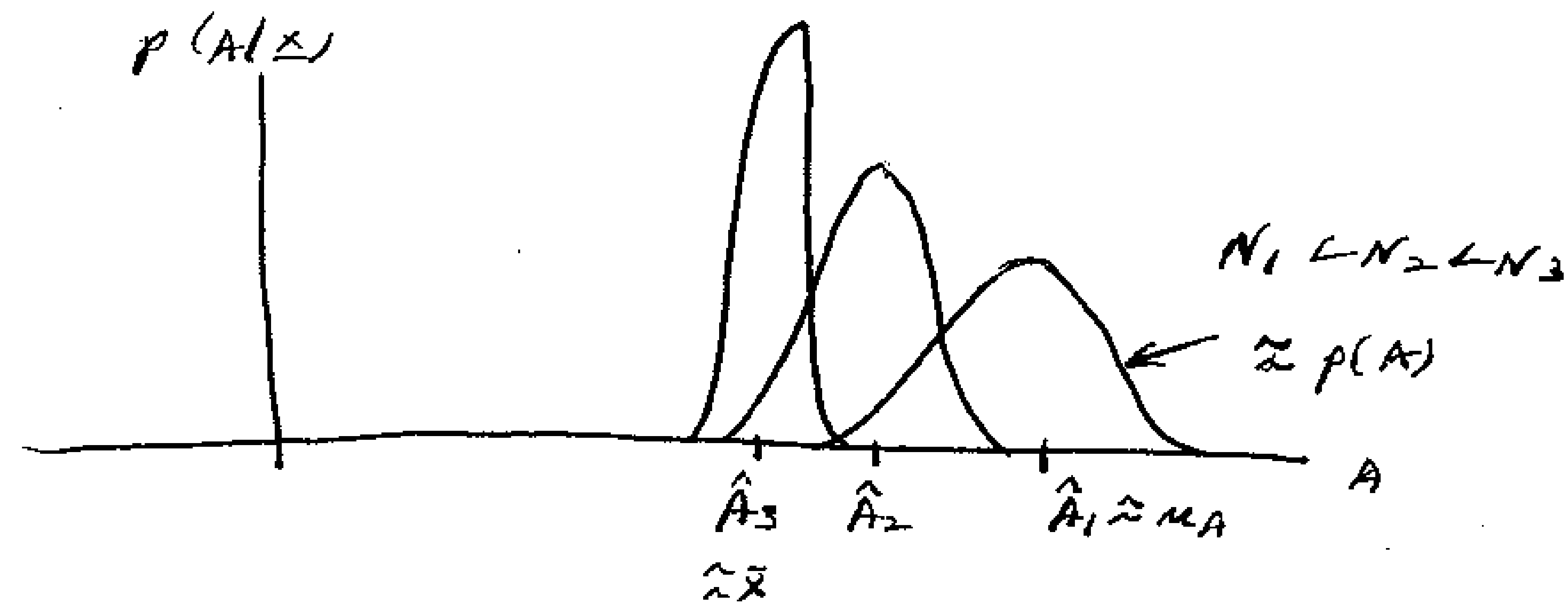
$$= \frac{\frac{N}{\sigma^2}\bar{x} + \frac{\mu_A}{\sigma_A^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}}$$

$$\hat{A} = \underbrace{\frac{\sigma_A^2}{\sigma_A^2 + \sigma^2/N}}_{\alpha} \bar{x} + \underbrace{\frac{\sigma^2/N}{\sigma_A^2 + \sigma^2/N}}_{1-\alpha} \mu_A$$

NOTE $\alpha$ IS WEIGHTING FACTOR $0 < \alpha < 1$, COMPROMISE BETWEEN $\mu_A$ (PRIOR ESTIMATE) AND $\bar{x}$ (DATA ESTIMATE).

FOR $\sigma_A^2 \ll \sigma^2/N \Rightarrow \alpha \approx 0 \Rightarrow \hat{A} \approx \mu_A$
FOR $\sigma_A^2 \gg \sigma^2/N \Rightarrow \alpha \approx 1 \Rightarrow \hat{A} \approx \bar{x}$

ALSO AS $N \to \infty$, $\sigma_A^2 \gg \sigma^2/N$ AND WE GET "CLASSICAL" ESTIMATOR $\bar{x}$.




TO SHOW THAT PRIOR KNOWLEDGE IMPROVES ESTIMATION ACCURACY AS MEASURED BY BAYESIAN MSE :

$$BMSE(\hat{A}) = E((\hat{A} - A)^2)$$

$$= \iint (\hat{A}-A)^2 \, p(\underline{x}, A) \, d\underline{x} \, dA$$

$$= \iint (\hat{A}-A)^2 \, p(A|\underline{x}) \, dA \, p(\underline{x}) \, d\underline{x}$$

$$= \int \underbrace{\int (A - E(A|\underline{x}))^2 \, p(A|\underline{x}) \, dA}_{VAR(A|\underline{x})} \, p(\underline{x}) \, d\underline{x}$$

$$BMSE(\hat{A}) = \int VAR(A|\underline{x}) \, p(\underline{x}) \, d\underline{x}$$

= VARIANCE OF POSTERIOR PDF AVERAGED OVER ALL REALIZATIONS OF $\underline{x}$.

$$BMSE(\hat{A}) = \int \sigma_{A|x}^2 \, p(\underline{x}) \, d\underline{x}$$

$$= \frac{1}{\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}}$$

$$= \frac{\sigma^2}{N} \left( \frac{\sigma_A^2}{\sigma_A^2 + \sigma^2/N} \right) < \sigma^2/N$$

$\sigma^2/N = BMSE(\hat{A})$ WHEN $\sigma_A^2 \to \infty$ OR NO PRIOR KNOWLEDGE

THINGS WORKED OUT WELL SINCE

$$\left.\begin{array}{l} p(A) = \text{GAUSSIAN} \\ p(A|\underline{x}) = \text{GAUSSIAN} \end{array}\right\} \text{REPRODUCING PROPERTY}$$

# PHYSICAL JUSTIFICATION FOR GAUSSIAN PRIOR

PROBLEM : MEASURE DC VOLTAGE OF INACCURATE POWER SOURCE WITH INACCURATE VOLTMETER

SET POWER SOURCE FOR 10 VOLTS, DUE TO ERRORS WE GET $A \sim N(10, \sigma_A^2)$. $\sigma_A^2$ CONTROLS OUR CONFIDENCE IN SOURCE.

VOLTMETER ADDS ERROR OR NOISE TO A

$$\Rightarrow \quad x[n] = A + w[n] \qquad n = 0, 1, \ldots, N-1$$

VOLTMETER ERROR VARIANCE $= \sigma^2$

TAKE N MEASUREMENTS.

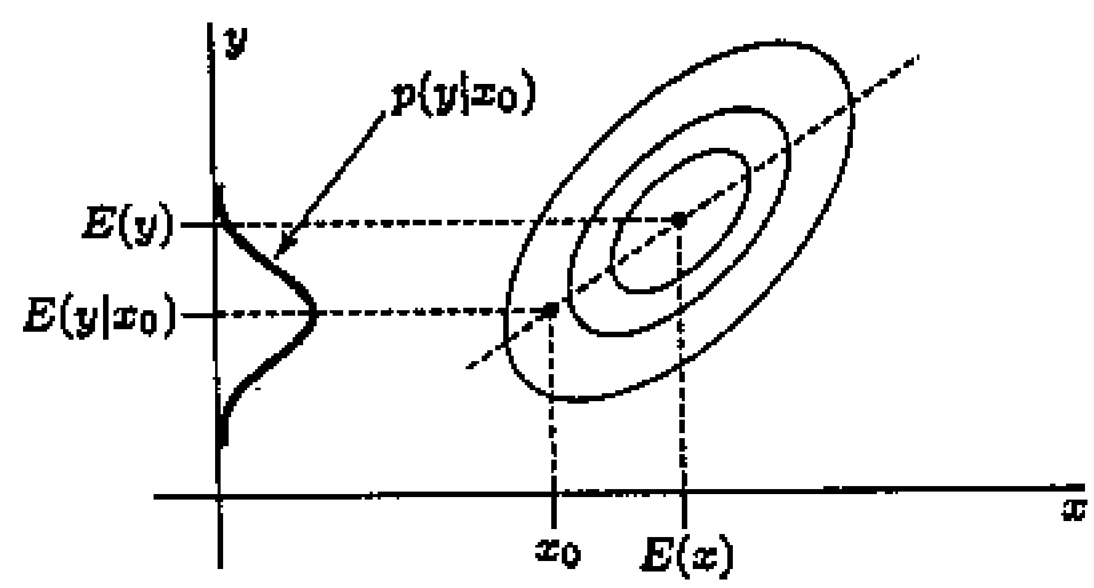
REPEAT PROCEDURE FOR ENSEMBLE OF POWER SOURCES ALL SET TO 10 VOLTS AND VOLTMETERS $\Rightarrow$ MSE MINIMIZED BY $\hat{A} = E(A \mid \underline{x})$.

# PROPERTIES OF GAUSSIAN PDF

$$p(x,y) = \frac{1}{2\pi |\underline{C}|^{\frac{1}{2}}} e^{-\frac{1}{2}\begin{bmatrix} x-E(x) \\ y-E(y)\end{bmatrix}^T \underline{C}^{-1} \begin{bmatrix} x-E(x) \\ y-E(y)\end{bmatrix}}$$

$$\underline{C} = \begin{bmatrix} VAR(x) & COV(x,y) \\ COV(y,x) & VAR(y) \end{bmatrix}$$

BY INTEGRATING $p(x,y)$ OVER $x$ OR $y$ WE FIND $x \sim N(E(x), VAR(x))$, $y \sim N(E(y), VAR(y))$.




CONTOURS OF CONSTANT PROBABILITY

$$\begin{bmatrix} x-E(x) \\ y-E(y)\end{bmatrix}^T \underline{C}^{-1} [\ \ ] = \text{CONSTANT}$$

$$p(y|x_0) = \frac{p(x_0,y)}{p(x_0)} = \frac{p(x_0,y)}{\int_{-\infty}^{\infty} p(x_0,y)\,dy}$$

CONDITIONAL PDF IS CROSS-SECTION OF $p(x,y)$ WHEN SUITABLY NORMALIZED TO INTEGRATE TO ONE. ALSO, $p(y|x_0)$ MUST BE GAUSSIAN SINCE $p(x_0,y)$ IS GAUSSIAN IN $y$.

$$E(y/x) = E(y) + \frac{COV(x,y)}{VAR(x)}(x - E(x))$$

$$VAR(y|x) = VAR(y) - \frac{COV^2(x,y)}{VAR(x)}$$

BEFORE OBSERVING $x$, $y \sim N(E(y), VAR(y))$
AFTER OBSERVING $x$, MEAN CHANGES AND VARIANCE DECREASES – LESS UNCERTAINTY ABOUT $y$

$$VAR(y|x) = VAR(y)\left[1 - \frac{COV^2(x,y)}{VAR(x)\,VAR(y)}\right]$$

$$= VAR(y)(1-\rho^2)$$

WHERE $\rho = \dfrac{COV(x,y)}{\sqrt{VAR(x)\,VAR(y)}}$ = CORRELATION COEFFICIENT

$$|\rho| \le 1$$

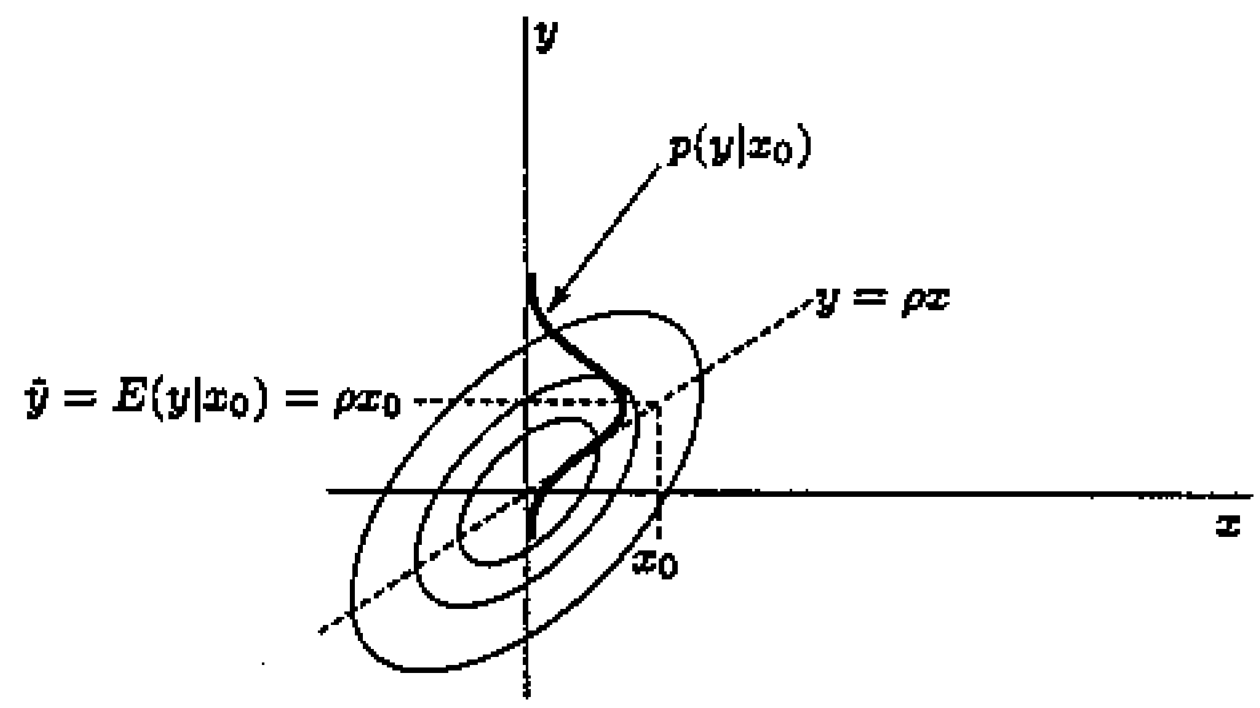
RECALL $E(y|x)$ IS MMSE ESTIMATOR OF $y$

$$\Rightarrow \quad \hat{y} = E(y) + \frac{COV(x,y)}{VAR(x)}(x - E(x))$$

OR

$$\underbrace{\frac{\hat{y} - E(y)}{\sqrt{VAR(y)}}}_{\hat{y}_n} = \underbrace{\frac{COV(x,y)}{\sqrt{VAR(x)\,VAR(y)}}}_{\rho}\ \underbrace{\frac{(x - E(x))}{\sqrt{VAR(x)}}}_{x_n}$$

↑ DENOTES NORMALIZED




EXPLOITS CORRELATION BETWEEN RANDOM VARIABLES

ALSO, $BMSE(\hat{y}) = \int VAR(y|x)\, p(x)\, dx$

$$= VAR(y|x) = VAR(y)(1-\rho^2)$$

IF $\rho = \pm 1$, WE ESTIMATE $y$ PERFECTLY
IF $\rho = 0$, $x, y$ ARE UNCORRELATED $\Rightarrow$ INDEPENDENT

$$\Rightarrow \hat{y} = E(y)$$
$$BMSE(\hat{y}) = VAR(y).$$

GENERALIZING TO MULTIVARIATE GAUSSIAN

$\underline{x}, \underline{y}$ JOINTLY GAUSSIAN

$\uparrow k\times 1$, $\uparrow l\times 1$

$$p(\underline{x}, \underline{y}) = \frac{1}{(2\pi)^{\frac{k+l}{2}} |\underline{C}|^{\frac{1}{2}}} e^{-\frac{1}{2}\begin{pmatrix}\underline{x}-E(\underline{x})\\ \underline{y}-E(\underline{y})\end{pmatrix}^T \underline{C}^{-1}\left[\quad\right]}$$

$\Rightarrow p(\underline{y}|\underline{x})$ ALSO MULTIVARIATE GAUSSIAN

$$E(\underline{y}|\underline{x}) = E(\underline{y}) + \underline{C}_{yx}\underline{C}_{xx}^{-1}(\underline{x}-E(\underline{x}))$$

$$\underline{C}_{y|x} = \underline{C}_{yy} - \underline{C}_{yx}\underline{C}_{xx}^{-1}\underline{C}_{xy}$$

WHERE $\underline{C} = \begin{bmatrix} \underline{C}_{xx} & \underline{C}_{xy} \\ \underline{C}_{yx} & \underline{C}_{yy} \end{bmatrix} = \begin{bmatrix} k \times k & k \times l \\ l \times k & l \times l \end{bmatrix}$

## BAYESIAN LINEAR MODEL

IN INTRODUCTORY EXAMPLE

$$x[n] = \underset{\mathcal{N}(\mu_A, \sigma_A^2)}{A} + w[n] \qquad n = 0, 1, \ldots, N-1$$

$w[n]$: WGN – INDEPENDENT OF $A$

$\Rightarrow \underline{x} = \underline{1} A + \underline{w}$

LOOKS LIKE LINEAR MODEL EXCEPT FOR A BEING RANDOM VARIABLE.

CONSIDER

$$\underline{x} = \underline{H}\,\underline{\theta} + \underline{w}$$

$\underline{x}$: $N \times 1$

$\underline{H}$: KNOWN $N \times p$ MATRIX

$\underline{w}$: $N \times 1$ RANDOM VECTOR, $\underline{w} \sim \mathcal{N}(\underline{0}, \underline{C}_w)$

NOW ASSUME $\underline{\theta}$ IS <u>RANDOM</u> <u>VECTOR</u> WITH PDF $\mathcal{N}(\underline{\mu}_\theta, \underline{C}_\theta)$ AND INDEPENDENT OF $\underline{w}$

$\Rightarrow$ <u>BAYESIAN</u> LINEAR MODEL.

NEED $p(\underline{\theta} | \underline{x})$ FOR BAYESIAN ESTIMATOR.