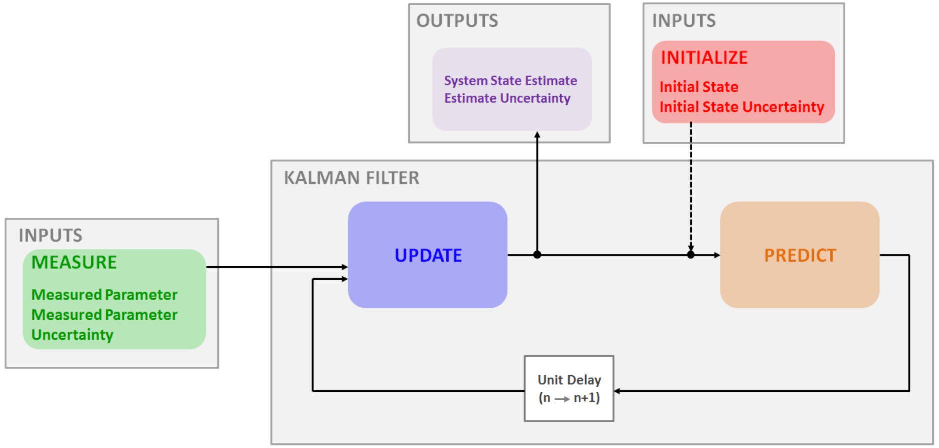
DATA FUSION: PROBABILISTIC AND DS BASED

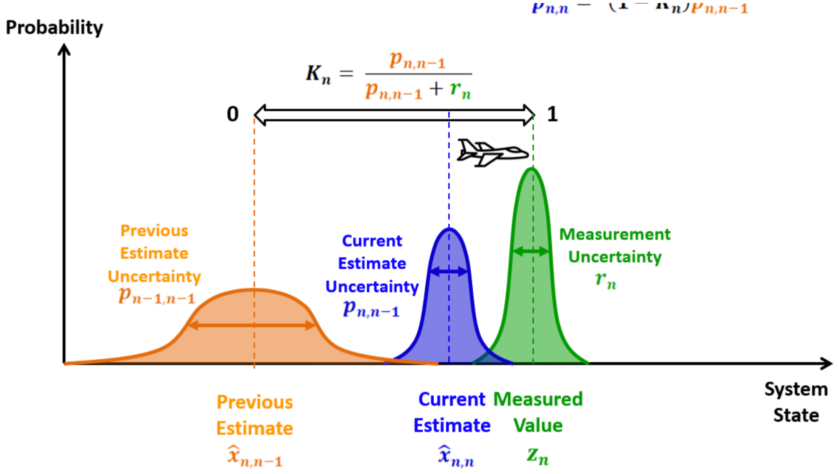
KALMAN FILTER DATA FUSION

The three KF widely used methods to perform fusion at the kinematic level are:

- (1) fusion of the raw observational and measurement data (the data con- verted to engineering units), called centralized fusion;
- (2) fusion of the estimated state vectors or state-vector fusion; and
- (3) the hybrid approach, which allows fusion of raw data and the processed state vector, as desired.



Kalman filtering has evolved to become a very high-level state-of-the-art method for estimation of the states of dynamic systems . The main reason for its success is that it has a very intuitively appealing state-space formulation and a predictor-corrector estimation and recursive-filtering structure; furthermore, it can be easily implemented on digital computers and digital signal processing units. It is a numerical data processing algorithm, which has tremendous real-time and online application potential.



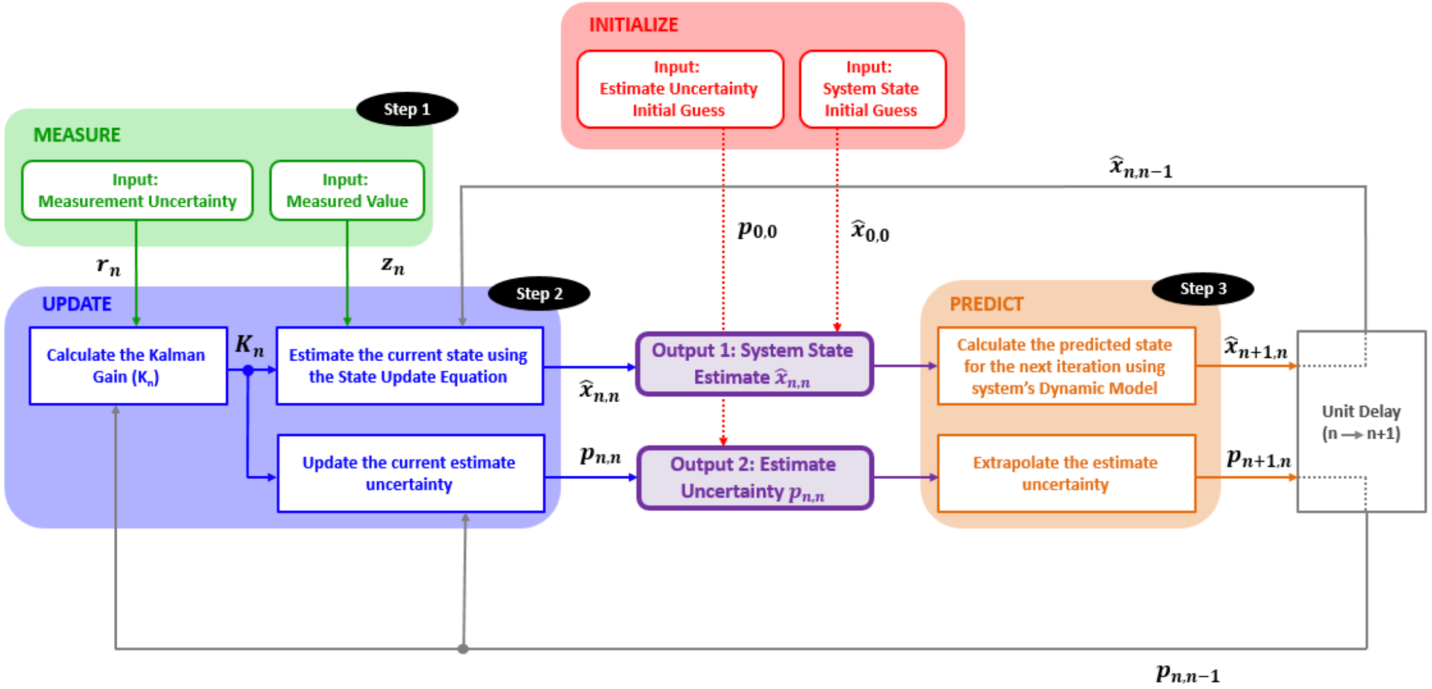
SIMPLE LOGIC

Recursive formulation:

new estimate = previous estimate + gain times the residuals of the estimation.

Very powerful and yet simple estimation data-processing structure.

Most of the real-time and online estimation and filtering algorithms have similar data-processing structures or algorithms.



KF FUSION

Here, x is the $n \cdot 1$ state vector;
u is the $px1$ control input vector to the dynamic system;
z is the $m \cdot 1$ measurement vector;
w is a white Gaussian process-noise sequence, with zero mean and covariance matrix Q;
v is a white Gaussian measurement-noise sequence, with zero mean and covariance matrix R;
ϕ is the $n \cdot n$ transition matrix that propagates the state (x) from k to $k+1$;
B is the input gain or magnitude vector or matrix;
H is the $m \cdot n$ measurement model or sensor-dynamic matrix; and
D is the $m \cdot p$ feed forward or direct-control input matrix, which is often excluded from the KF development.
In addition, B is often omitted if there is no explicit control input playing a role.

Dynamic system model:

$$x(k+1) = \phi x(k) + Bu(k) + Gw(k) \tag{2.17}$$

$$z(k) = Hx(k) + Du(k) + v(k) \tag{2.18}$$

Modification of the KF with inclusion of B and D is relatively straightforward.

Although most dynamic systems are continuous in time, the KF technique is the best discussed and is mostly used in the discrete-time form.

The problem of state estimation using KF is formulated as follows:

given the model of the dynamic system, statistics regarding the noise (Q, R) processes, the noisy measurement data (z), and the input (u), determine the optimal estimate of the state, x, of the system.

We presume that the state estimate at k has evolved to $k+1$. At this stage, a new measurement is made available, and it hopefully contains new information regarding the state, as per Equation 2.18.

Hence, the idea is to incorporate the measurement into the data-fusion (i.e., update or filtering) process and obtain an improved and refined estimate of the state.

DATA UPDATE ALGORITHM

We have the measurement z, know H, and have assumed R; we further have/assume $\tilde{x}(k) \rightarrow$ the "a priori" estimate of state at time k, i.e., before the measurement data is incorporated, and $\tilde{P} \rightarrow$ the "a priori" covariance matrix of the state-estimation error (the time-index k is omitted for simplifying the equations). Then, the measurement-update algorithm (essentially, the filtering of the state vector by considering the measurement

data) to obtain $\hat{x}(k) \rightarrow$ the updated estimate of state at time k, i.e., after the measurement data is incorporated, is given as:

Residual equation:

$$r(k) = z(k) - H\,\tilde{x}(k) \tag{2.19}$$

Kalman gain:

$$K = \tilde{P}H^T(H\tilde{P}H^T + R)^{-1} \tag{2.20}$$

Filtered state estimate:

$$\hat{x}(k) = \tilde{x}(k) + K\,r(k) \tag{2.21}$$

Covariance matrix (posteriori):

$$\hat{P} = (I - KH)\tilde{P} \tag{2.22}$$

STATE PROPAGATION

This part of the KF method, which applies the previous estimates of x and P, is represented as:

State estimate:

$$\tilde{x}(k+1) = \phi\ \hat{x}(k) \tag{2.23}$$

Covariance matrix:

$$\tilde{P}(k+1) = \phi \hat{P}(k) \phi^T + GQG^T \tag{2.24}$$

In the KF method, $K = \tilde{P}H^T S^{-1}$ and $S = H\tilde{P}H^T + R$, and matrix S is the covariance matrix of residuals (also called innovations). The actual residuals can be computed and compared with the standard deviations obtained by calculating the square root of the diagonal elements of S. The process of tuning the filter to bring the computed residuals within the bounds of at least two standard deviations is an important filter-tuning exercise for obtaining the correct solution to the problem [14].

This is the measurement data–level fusion algorithm. It combines the measurements of these observables directly at the data level and produces an optimal estimate of the state x.

For each measurement type, one should choose appropriate H vectors or matrices and their corresponding R matrices. Therefore, the KF fundamentally accomplishes a DF task.

BAYESIAN AND DEMPSTER-SHAFER FUSION

- KF can be viewed as a prediction−corrector (state propagation or evolution and data updating) filtering algorithm, which is widely used for tracking moving object and targets.
- In addition, the KF is in itself a data–level fusion algorithm. It is possible to consider KF as a Bayesian fusion algorithm.
- The Bayesian approach involves the definition of priors (*a priori* probabilities), their specifications, and computations of the posteriors.
- Primarily, the probability theory is based on crisp logic, comprising "zero" or "one" (yes or no, on or off, -1 or $+1$).
- It does not consider any third possibility, because the probability definition is based on set theory, which is in turn based on crisp logic; it considers only the probability of occurrence or non−occurrence of an event.

BAYESIAN METHOD

Let us define $p(A)$ as the probability of occurrence of an event A, and $p(A, B)$ as the probability of occurrence of two events A and B. Then, the conditional probability of occurrence of A, given that the event B has already occurred, can be related as follows:

$$p(A,\ B) = p(A\,|\,B)\ p(B)$$

We will also notice that because $p(A, B) = p(B, A)$,

$$p(B,\ A) = p(B\,|\,A)\ p(A)$$

$$p(A\,|\,B) = p(B\,|\,A)\ p(A)/p(B)$$

The above relation can also be written as follows:

$$p(A\,|\,B) = \frac{p(B\,|\,A)\,p(A)}{\sum_i p(B\,|\,A_i)\,p(A_i)} \tag{2.25}$$

The denominator acts as a normalization factor if there are several events of A that can be distinguished from B in a few ways. The above equation is known as Bayes' rule. Replacing A with x and B with z, we obtain the following relation:

$$p(x\,|\,z) = \frac{p(z\,|\,x)\,p(x)}{\sum_i p(z\,|\,x_i)\,p(x_i)} \tag{2.26}$$

BAYES FUSION

The items x and z are regarded as random variables, and x is a state or parameter of the system and z is the sensor measurements.

Thus, the Bayes' theorem is interpreted as the computation of the posterior probability, given the prior probability of the state or parameter $(p(x))$, and the observation probability $(p(z|x))$: the value of x that maximizes the term $(x|data)$.

The maximum likelihood is related to this term if $p(x)$ is considered a uniform distribution: the value of x that maximizes the term $(data|x)$.

The term $p(z|x)$ assumes the role of a sensor model in the following way:

(1) First build a sensor model: fix x, and then find the probability- density function (pdf) on z.

(2) Use the sensor model: observe z, and then find the pdf on x.

(3) For each fixed value of x, a distribution in z is defined.

(4) As x varies, a family of distributions in z is formulated. For the observation z of a target-tracking problem with state x, the Gaussian-observation model is given as a function of both z and x as shown below:

BAYES FUSION

$$p(z \mid x) = \frac{1}{\sqrt{2\pi}\ \sigma_z^2} \exp\left(-\frac{(z-x)^2}{2\sigma_z^2}\right) \tag{2.27}$$

When the model is built, the state is fixed, and the distribution is then a function of z. When the observations are made, the distribution is a function of x. The prior $p(x)$ is given as

$$p(x) = \frac{1}{\sqrt{2\pi}\ \sigma_x^2} \exp\left(-\frac{(x-x_p)^2}{2\sigma_x^2}\right) \tag{2.28}$$

Then, using the Bayes' rule, the posterior is given as follows after noting the observation:

$$\begin{aligned} P(x/z) &= \text{Const.}\frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left(-\frac{(z-x)^2}{2\sigma_z^2}\right)\frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left(-\frac{(x-x_p)^2}{2\sigma_x^2}\right) \\ &= \frac{1}{\sqrt{2\pi}\ \sigma^2} \exp\left(-\frac{(x-\overline{x})^2}{2\sigma^2}\right) \end{aligned} \tag{2.29}$$

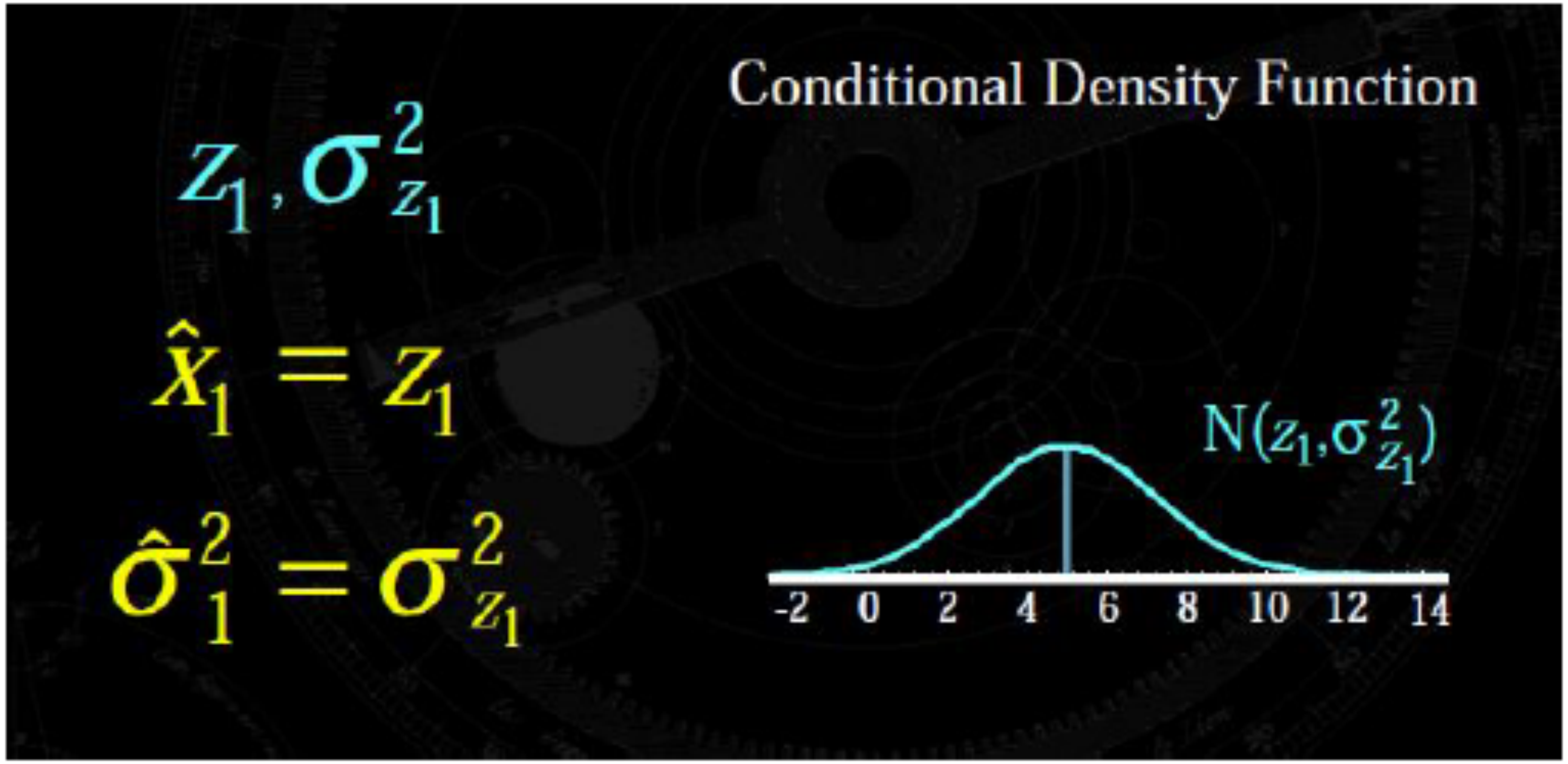
where

$$\overline{x} = \frac{\sigma_x^2}{\sigma_x^2+\sigma_z^2} z + \frac{\sigma_z^2}{\sigma_x^2+\sigma_z^2} x_p \tag{2.30}$$

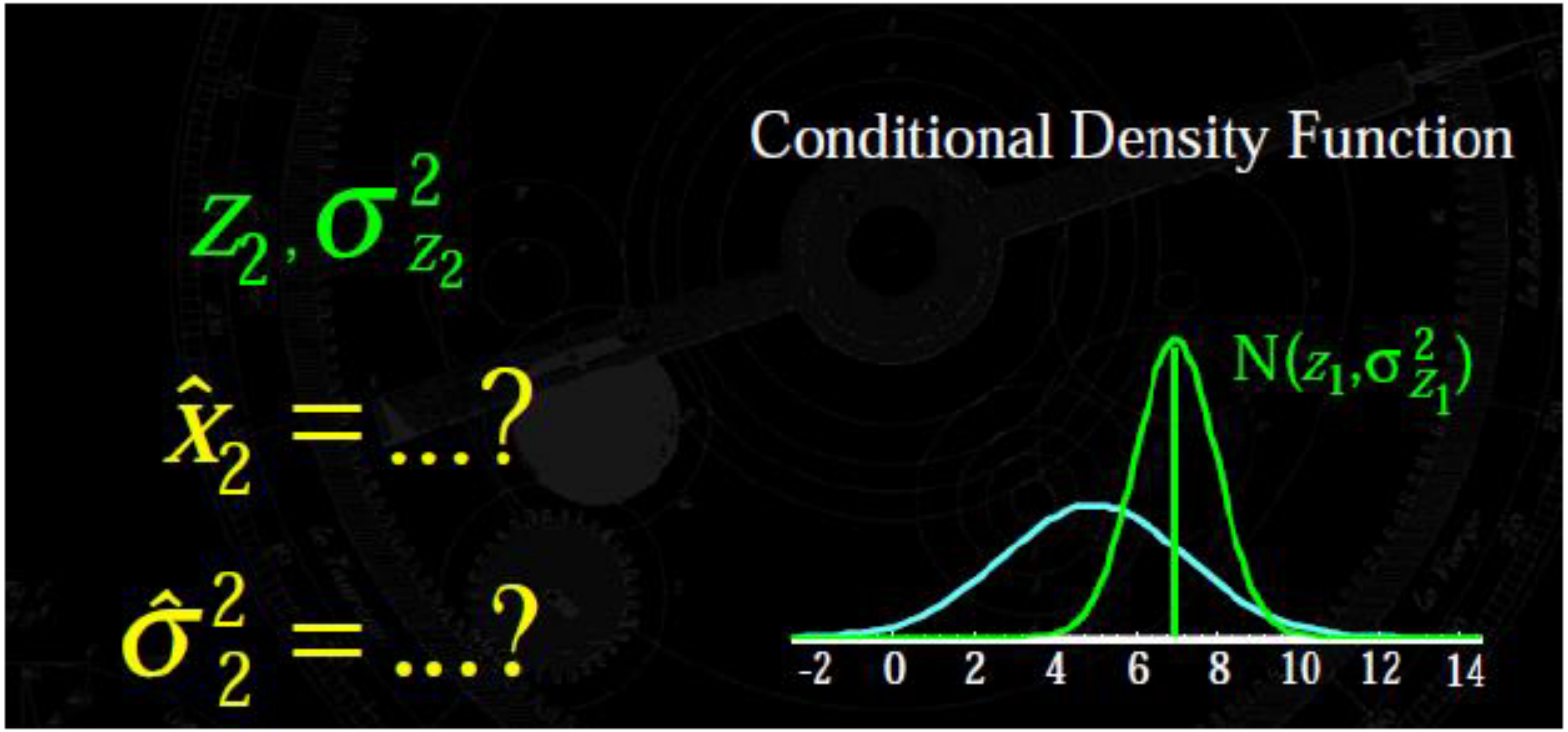
and

$$\sigma^2 = \frac{\sigma_z^2\sigma_x^2}{\sigma_x^2+\sigma_z^2} = \left(\frac{1}{\sigma_z^2}+\frac{1}{\sigma_x^2}\right)^{-1} \tag{2.31}$$

SEQUENTIAL DATA FUSION



NEW OBSERVATIO



COMBINING THE TWO OBSERVATIONS

$$\hat{x}_2 = \hat{x}_1 + K_2\left(z_2 - \hat{x}_1\right)$$

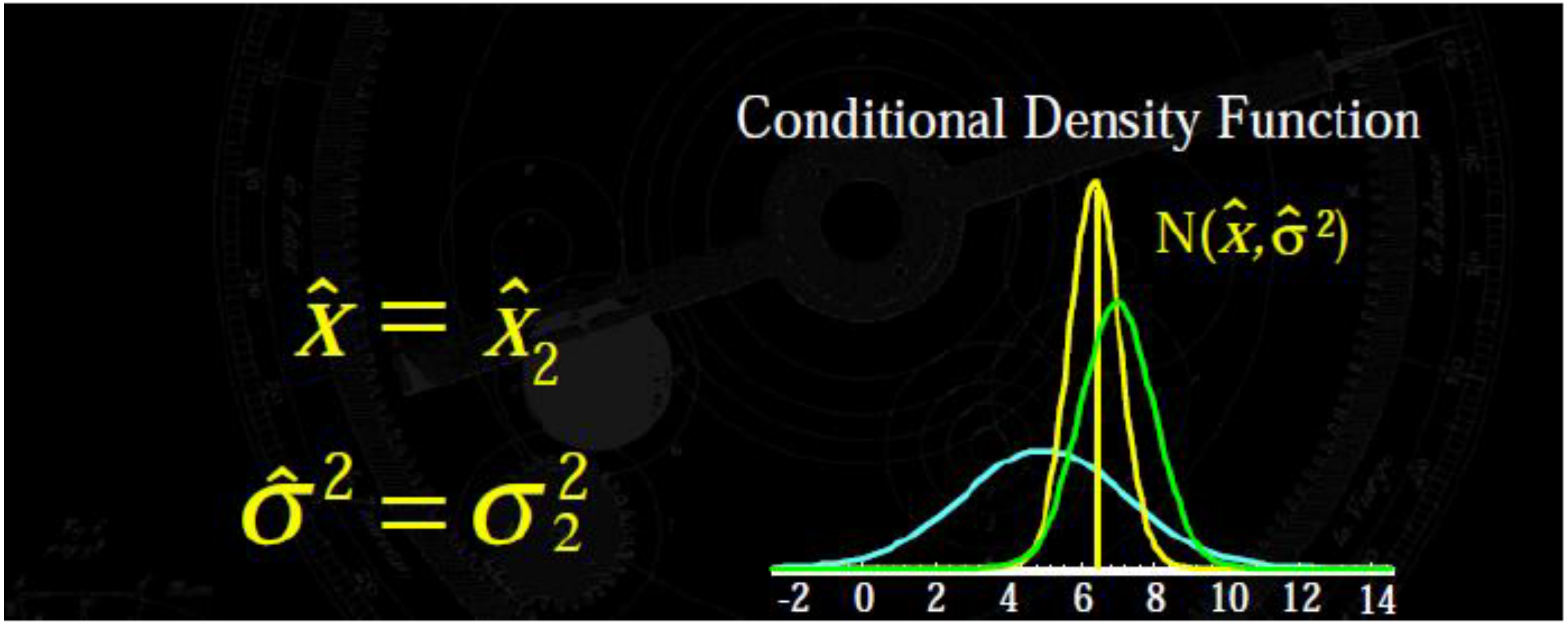
$$K_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_{z_2}^2}$$

LESS VARIATION

$$\frac{1}{\sigma_2^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_{z_2}^2}$$

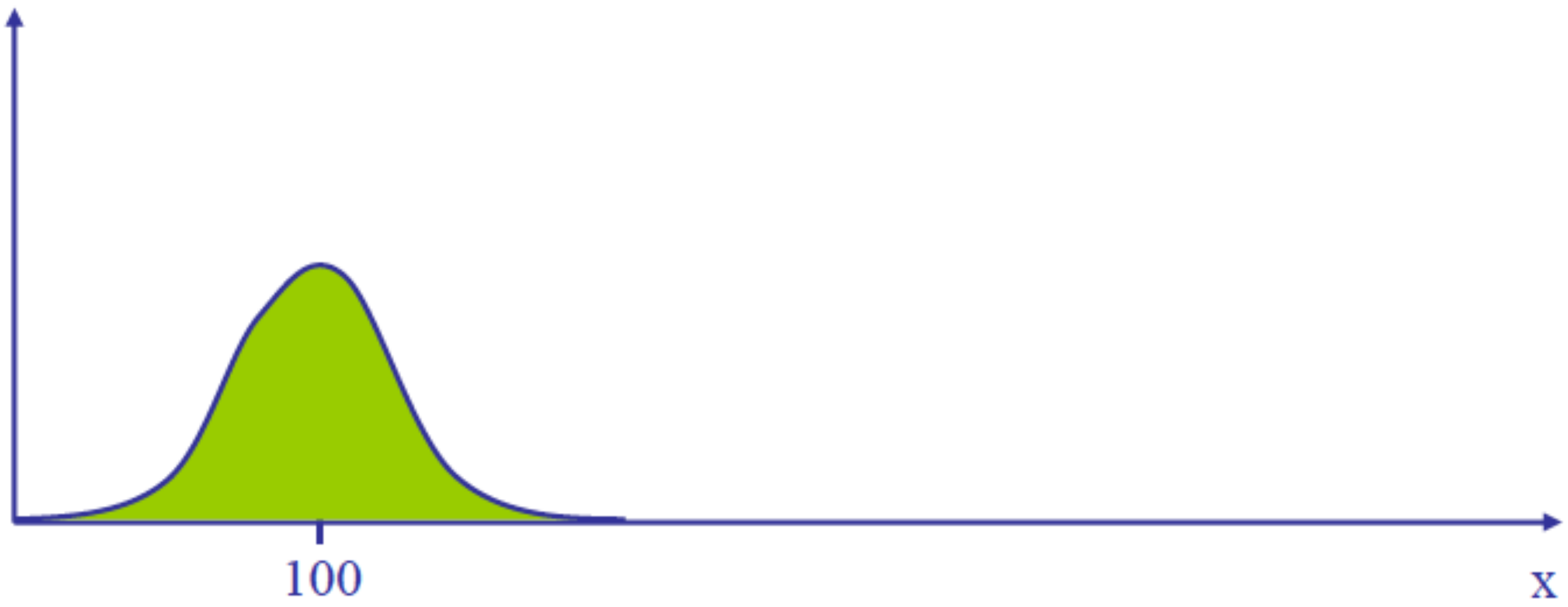
Obviously the variance of the combination is smaller than the (or at most equal to) the smallest variance

IMPROVED ACCURACY



EXAMPLE

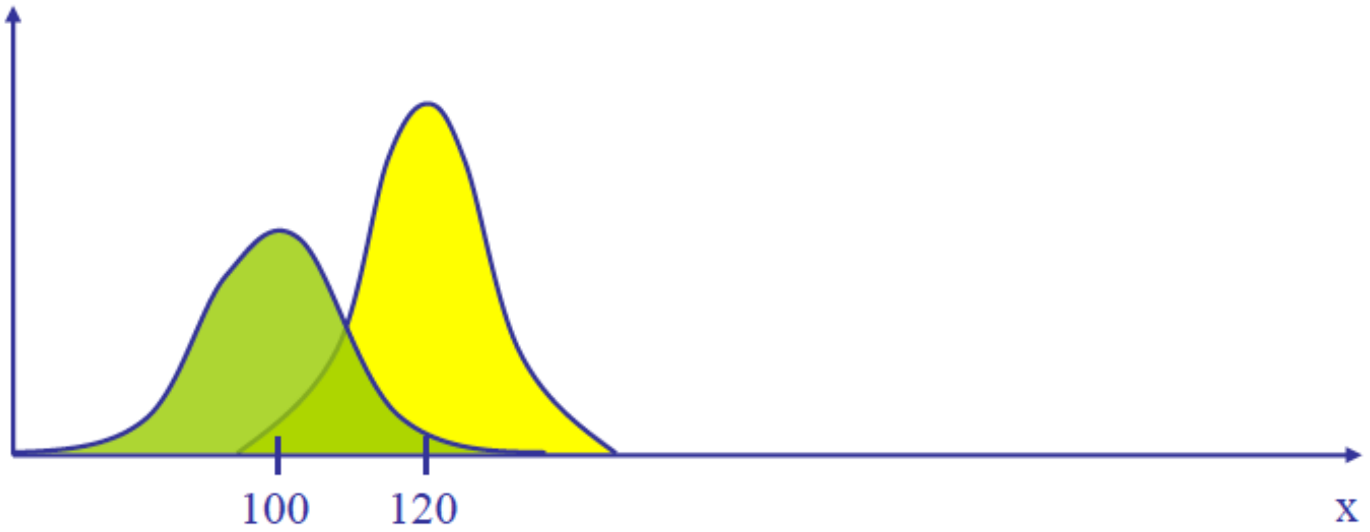
We estimate the position x from the start as Z1=100 with a std variation of 4 miles



NEW OBSERVATION

a more accurate sensor provides z_2

- With position estimate $x = z_2 = 125$ with a precision of $\sigma_x = 3$ miles
- combining?



$$\mu = \left[\frac{\sigma_{z_2}^2}{\sigma_{z1}^2 + \sigma_{z_2}^2}\right] z_1 + \left[\frac{\sigma_{z_1}^2}{\sigma_{z1}^2 + \sigma_{z_2}^2}\right] z_2$$

$$= \left[\frac{9}{16+9}\right] 100 + \left[\frac{16}{16+9}\right] 125 = 116$$

$$\frac{1}{\sigma^2} = \frac{1}{\sigma_{z_1}^2} + \frac{1}{\sigma_{z_2}^2}$$

$$\frac{1}{\sigma^2} = \frac{1}{9} + \frac{1}{16} = \frac{25}{144}$$

$$\Rightarrow \sigma = 2.4$$