# 9 Object Micro-Doppler Signature Parameter Learning

We now turn our attention to the problem of parameter learning in graphical models. There are various established approaches to this problem, ranging from optimisation methods for finding point estimates for parameters to more bayesian treatments. For our application we will focus on the bayesian view of parameter learning. The bayesian approach to parameter learning is to treat the parameters of the network as random variables and to then make inferences about the distribution of these variables as would be done for any other variables in the network. This approach therefore offers a bayesian framework for machine learning in that we can learn distributions over parameters from data. The bayesian approach offers a couple of advantages over other methods ( such as point estimate optimisation - typically used in many machine learning algorithms). These include the ability to include prior knowledge into the model which can greatly reduce the amount of data that is needed to train a model. These methods also often lead to better generalisation and more more accurate uncertainty measures in areas of the data space where few training examples are available. The bayesian approach does however also have disadvantages, these methods are more computationally expensive than optimisation based methods and can be intractable for large models. Furthermore, analytical or closed form solutions to the involved integrals are often not available, although approximate methods, such as variational methods have been developed and successfully used to address this issue.

In our specific application, we will be interested in learning distributions over variables that can help characterise and identify the objects that the model is tracking. The advantage of incorporating such functionality into the model will be twofold. First, it will improve tracking performance. Specifically, it will help to disambiguate many inherently ambiguous scenarios (that occur when two targets are close to each other and have similar velocities) as well as improve inference about the number of targets. Secondly, it will allow us to learn something about the characteristics of the targets, which can then help to perform inference about the similarity of targets and to classify the target type under certain circumstances [4].

No that we have a basic overview of the problem, the solution approach and the benefits of implementing the additional functionality, we will investigate on more detail how we might implement the necessary model. First, we will consider what data we have available and how we can use this data. The micro-doppler data that is a available as an output from the radar is an obvious starting point for this process. The micro-doppler data, contains more information (compared to the doppler data) about the velocities of moving objects in area bins. We might therefore expect this data to include some information about appendages moving relative to a target (a person with swinging arms for example). This could in theory be useful in forming a probabilistic signature for a target. At this point we must also notice that, although this data may be useful, we cannot use it directly to learn a characteristic distributions of targets. The reason for this is that a target appendages moving in a specific and constant way result in different micro-doppler data being generated as a result of a change in the velocity size or direction of the target. This is of course because the radar does not measure the micro-doppler velocities relative to the velocity of the object. We will therefore need to perform some random variable transformations in the network in order to obtain information about the relative (to the speed and direction of the target) movement

[4]Of course we will need to know the labels of the learnt characteristic distributions in order to actually classify new targets, although we can atleast perform clustering without labels

of target appendages and infer the characteristic signature distributions. These transformations are similar to the cartesian to polar transformations that have previously been discussed and used in the multiple target tracking model that was developed. Before discussing the specifics of these transformations, we will first construct a bayes network for the extended model in order lay down the basic structure of the model. The will help us to first ensure that the model makes logical sense and give some hints as to the types of transformations we will need. As a starting point for the extended model, we will use the Kalman Filter Bayes network for simplicity. We will however view this as a simplified version of the multiple object tracking Bayes network that was developed. Extending the model for multiple object tracking This should be straight forward as there will not be any edges between the any of the new nodes for different targets.

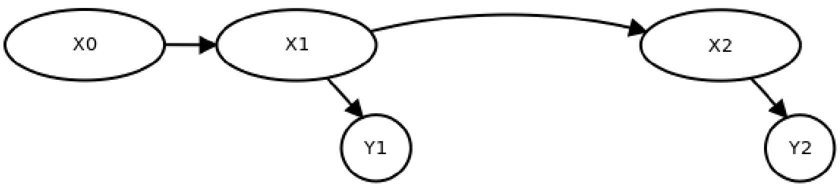


Figure 20: Kalman Filter Bayes Network

We start our design by noticing that the micro-doppler signal is similar to the doppler (radial velocity) variable that already exists in the existing network. We will therefore add a random variable (actually a random vector, as the micro-doppler data is multidimensional) node $U$ to the network to represent the cartesian transformation of the micro-doppler velocities. The actual micro-doppler variable that will be observed can simply be included in the $Y$ node, as this node already represents a random vector as it is used in our application. As we have discussed earlier however, we are not particularly interested in the cartesian transformed representation of these velocities either. What we are interested in are the micro-doppler velocities relative to the velocity of the target. We therefore need to add another type of node $S$ to represent these relative velocities. The $U$ velocities will therefore be dependent on the target's velocity $V$ and the relative velocities $S$. The additional nodes and edges are shown in the graph below. The $W$ and $V$ variables represent the cartesian position and velocity random vectors respectively and are included as a reminder that the $U$ variable is only dependent on $V$ in $X$.

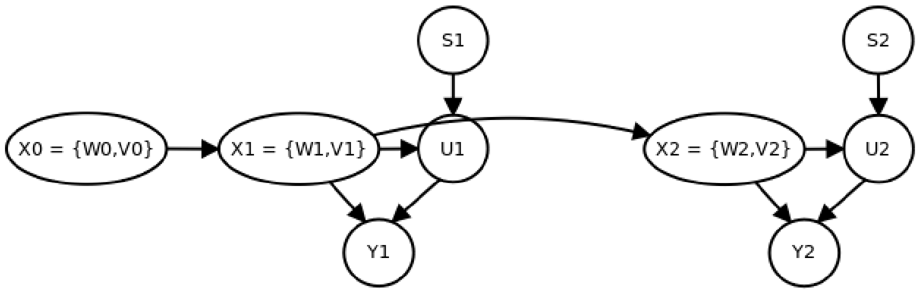


Figure 21: Bayes Network

We will now derive the necessary functions to transform between the various forms of the micro-doppler velocities, using the diagram below, which shows how the different representations relate to each other.

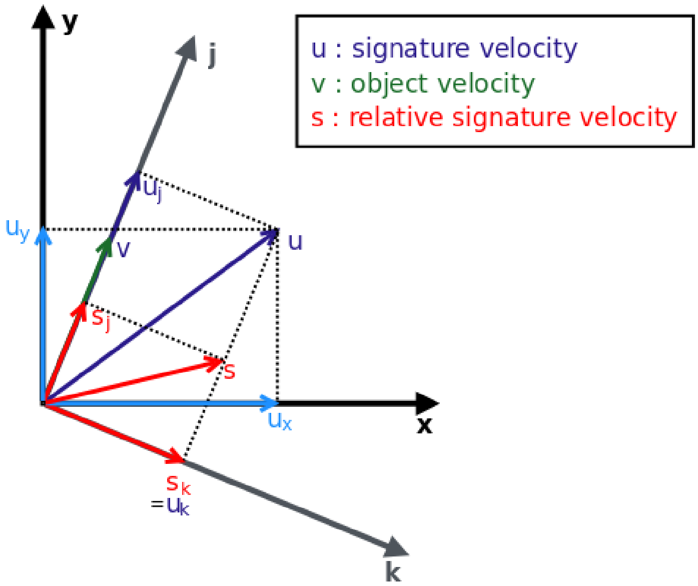


Figure 22: Object velocity and micro-doppler velocity relation diagram

From the above diagram we can see that

$$
\begin{aligned}
u_j =& s_j + v \\
u_k =& s_k
\end{aligned}
$$

And that

$$
\begin{aligned}
u_j =& \frac{u_x v_x + u_y v_y}{|v|} \\
u_k =& \frac{u_x v_y - u_y v_x}{|v|}
\end{aligned}
$$

Rewriting the above expressions in terms of $u_x$ and $u_y$, we get

$$u_x = \frac{u_j v_x + u_k v_y}{|v|}$$
$$u_y = \frac{u_j v_y - u_k v_x}{|v|}$$

And finally, in terms of $s$ and $v$

$$\begin{aligned} u_x &= \frac{(s_j + |v|)v_x + s_k v_y}{|v|} \\ &= \frac{(s_j + \sqrt{v_x^2 + v_y^2})v_x + s_k v_y}{\sqrt{v_x^2 + v_y^2}} \end{aligned}$$

$$\begin{aligned} u_y &= \frac{(s_j + |v|)v_y - s_k v_x}{|v|} \\ &= \frac{(s_j + \sqrt{v_x^2 + v_y^2})v_y - s_k v_x}{\sqrt{v_x^2 + v_y^2}} \end{aligned}$$

So now we have derived the necessary transforms for calculating $U$ from $V$ and $S$. For the transformation from $U$ to the polar micro-doppler components in $Y$, we can simply use the standard cartesian-polar sigma-point transformations as derived earlier. We can also use the sigma-point representation and transform with the $U$ transform functions above.

The network is however still not complete, as we still do not have any way of learning a distribution over $S$. In order to facilitate this functionality in the bayesian spirit, we will place priors over the covariance matrix and mean of $S$ and because we will make the assumption that the distribution over $S$ (for a certain target) will be constant over time, these prior nodes will be connected to all time instances of $S$. The full network, with parameter priors as well as the correspond two time-slice Bayes network (for the general multiple object tracking version) are shown in the following figures.

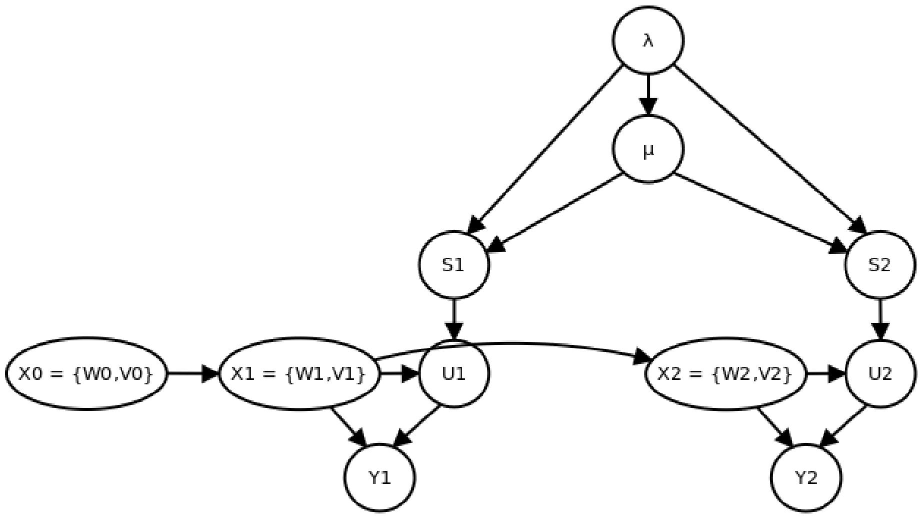


Figure 23: Bayes Network

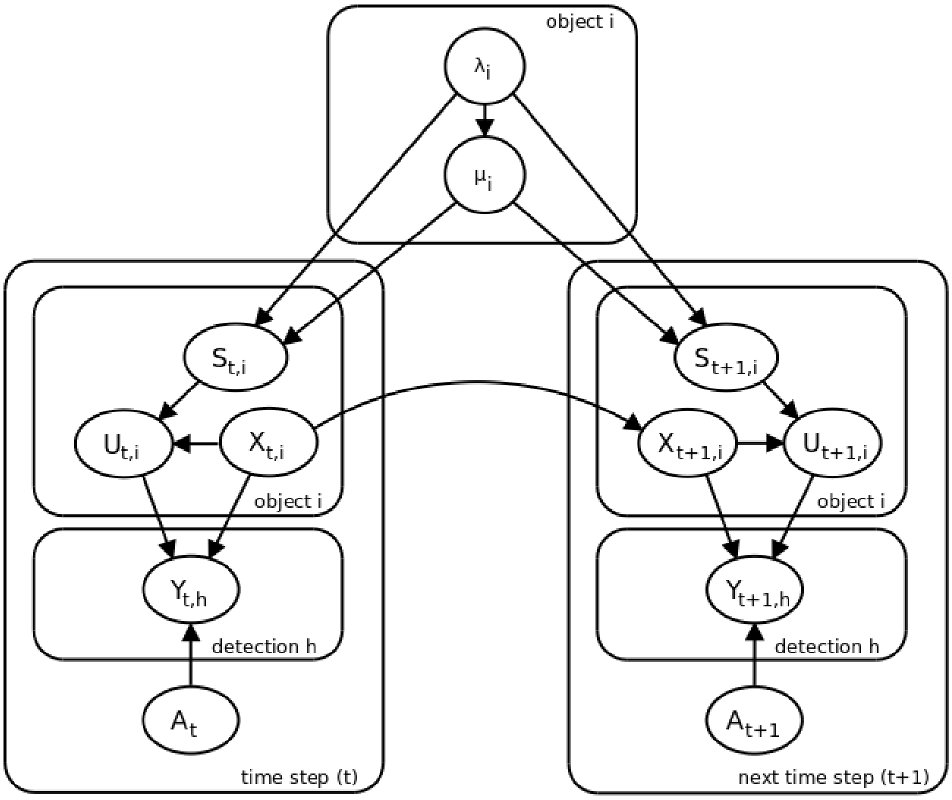


Figure 24: Two Time-Slice Bayes Network for the multiple object tracking and signature learning

In order to derive the corresponding cluster-graph, we will perform the normal process of moralisation, triangulation and clustering. The moralised graph is shown below.

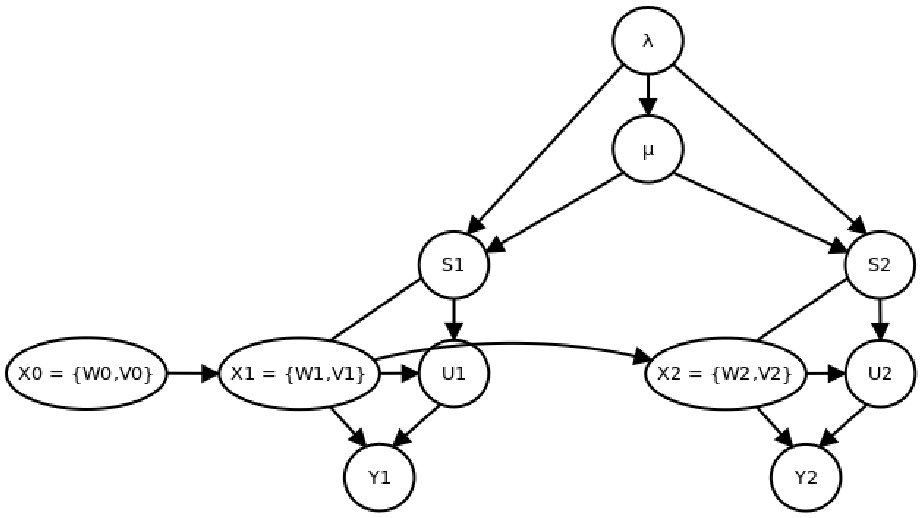


Figure 25: Moralised Bayes Network 16

The inducing markov network, obtained by replacing the directed edges with undirected edges, is shown below.

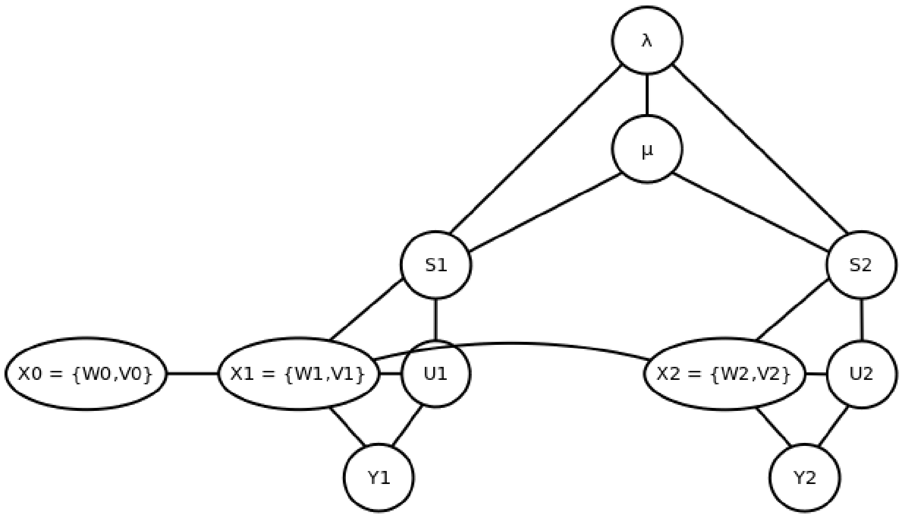


Figure 26: Induced Markov Network

And finally, by clustering variables each clique in the induced markov network, we get the cluster graph shown below.

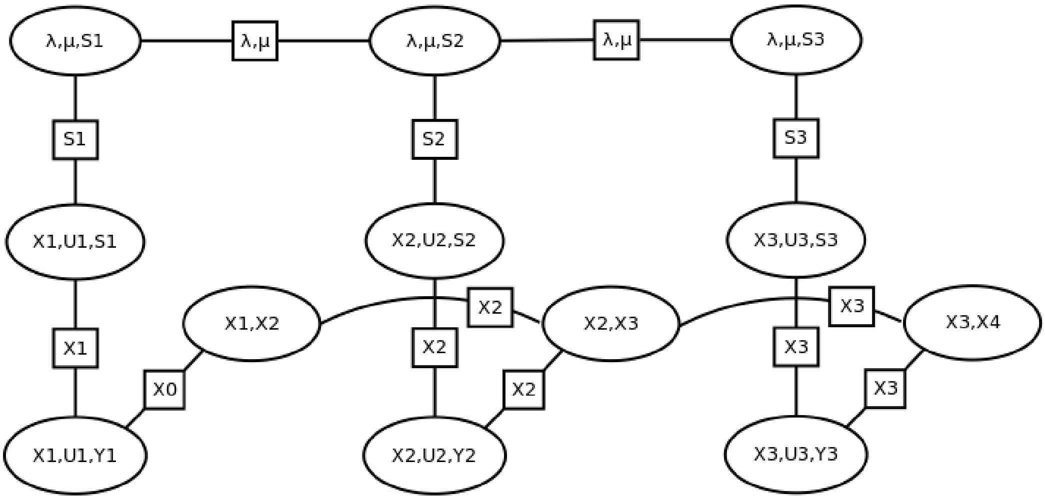


Figure 27: Micro-doppler signature learning and object tracking Cluster Graph

Now that we have the variable transformation functions and the structure of the bayes network and cluster graph, we still need to define the forms of the $P(\mu|\lambda)$ and $P(\lambda)$ (or $P(\mu|K)$ and $P(K)$ in the multivariate case) prior distributions.

## 10 The Normal-Wishart Distribution

$$P(K) = \frac{|K|^{(v-d-1)/2}}{2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)} e^{-tr(V^{-1}K)/2}$$

$$P(\mu|K) = \frac{|\lambda_0 K|^{1/2}}{(2\pi)^{d/2}} e^{-0.5([\mu-\mu_0]^T \lambda_0 K[\mu-\mu_0])}$$

$$P(X|\mu, K) = \frac{|K|^{1/2}}{(2\pi)^{d/2}} e^{-0.5([X-\mu]^T K[X-\mu])}$$

where $K$ is the precision matrix of the gaussian distribution over $X$. Note that, in contrast to previous use of the gaussian expression, the $K$ matrix and $\mu$ vector in the above expression are not constants but a random vector and matrix respectively. The above distribution is therefore no longer gaussian, but rather a conditional distribution over gaussian distributions, we're any observation of the mean and precision will collapse the distribution into a gaussian distribution. Luckily, it seems that we will only every have to marginalise out $X$ or $K$ and $\mu$ during inference.

$$P(X, \mu, K) = P(X|\mu, K)P(\mu|K)P(K)$$
$$P(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(X|\mu, K)P(\mu|K)P(K)d\mu dK$$

$$\begin{aligned}
P(\mu, K) &= \int_{-\infty}^{\infty} P(X|\mu, K)P(\mu|K)P(K)dX \\
&= \int P(X|\mu, K)dX P(\mu|K)P(K) \\
&= P(\mu|K)P(K)
\end{aligned}$$

$$\begin{aligned}
P(X) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(X|\mu, K)P(\mu|K)P(K)d\mu dK \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{|K|^{(v-d-1)/2}}{2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)} \frac{|\lambda_0 K|^{1/2}}{(2\pi)^{d/2}} \frac{|K|^{1/2}}{(2\pi)^{d/2}} e^{F(X,\mu,K)} d\mu dK \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\lambda_0^{d/2}|K|^{(v-d+1)/2}}{2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^d} e^{-0.5F(X,\mu,K)} d\mu dK
\end{aligned}$$

$$
\begin{aligned}
F(X,\mu,K) =& tr(V^{-1}K) \\
& + ([\mu-\mu_0]^T\lambda_0 K[\mu-\mu_0]) \\
& + ([X-\mu]^T K[X-\mu]) \\
=& tr(V^{-1}K) \\
& + \mu^T\lambda_0 K\mu + \mu_0^T\lambda_0 K\mu_0 - 2\mu^T\lambda_0 K\mu_0 \\
& + X^T KX + \mu^T K\mu - 2X^T K\mu \\
=& tr(V^{-1}K) \\
& + \mu^T\lambda_0 K\mu + \mu^T K\mu - 2\mu^T KX - 2\mu^T\lambda_0 K\mu_0 \\
& + X^T KX + \mu_0^T\lambda_0 K\mu_0 \\
=& tr(V^{-1}K) \\
& + \mu^T(\lambda_0+1)K\mu - 2\mu^T K(X+\lambda_0\mu_0) \\
& + X^T KX + \mu_0^T\lambda_0 K\mu_0 \\
=& tr(V^{-1}K) \\
& (\mu-\omega)^T(\lambda_0+1)K(\mu-\omega) - \omega^T(\lambda_0+1)K\omega \\
& + X^T KX + \mu_0^T\lambda_0 K\mu_0 + \\
=& tr(V^{-1}K) \\
& (\mu-\omega)^T(\lambda_0+1)K(\mu-\omega) \\
& + X^T KX + \mu_0^T\lambda_0 K\mu_0 - \omega^T(\lambda_0+1)K\omega \\
=& (\mu-\omega)^T(\lambda_0+1)K(\mu-\omega) \\
& + tr(V^{-1}K) + tr(X^T KX) + tr(\mu_0^T\lambda_0 K\mu_0) - tr(\omega^T(\lambda_0+1)K\omega) \\
=& (\mu-\omega)^T(\lambda_0+1)K(\mu-\omega) \\
& + tr(V^{-1}K + XX^T K + \mu_0\mu_0^T\lambda_0 K - \omega\omega^T(\lambda_0+1)K) \\
=& (\mu-\omega)^T(\lambda_0+1)K(\mu-\omega) + H(X,K)
\end{aligned}
$$

with

$$\omega = \frac{X+\lambda_0\mu_0}{1+\lambda_0}$$

$$
\begin{aligned}
H(X,K) =& tr(V^{-1}K + XX^T K + \mu_0\mu_0^T\lambda_0 K - \omega\omega^T(\lambda_0+1)K) \\
=& tr((V^{-1} + XX^T + \lambda_0\mu_0\mu_0^T - (\lambda_0+1)\omega\omega^T)K)
\end{aligned}
$$

So, we can rewrite the integral as:

$$
\begin{aligned}
P(X) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\lambda_0^{d/2}|K|^{(v-d+1)/2}}{2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^d} e^{-0.5F(X,\mu,K)} d\mu dK \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\lambda_0^{d/2}|K|^{(v-d+1)/2}}{2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^d} e^{-0.5((\mu-\omega)^T(\lambda_0+1)K(\mu-\omega)+H(X,K))} d\mu dK \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\lambda_0^{d/2}|K|^{(v-d+1)/2}}{2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^d} e^{-0.5(\mu-\omega)^T(\lambda_0+1)K(\mu-\omega)} e^{-0.5H(X,K))} d\mu dK \\
&= \int_{-\infty}^{\infty} \frac{\lambda_0^{d/2}|K|^{(v-d+1)/2}}{2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^d} \left(\int_{-\infty}^{\infty} e^{-0.5(\mu-\omega)^T(\lambda_0+1)K(\mu-\omega)} d\mu\right) e^{-0.5H(X,K))} dK \\
&= \int_{-\infty}^{\infty} \frac{\lambda_0^{d/2}|K|^{(v-d+1)/2}}{2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^d} \left(\frac{(2\pi)^{d/2}}{(\lambda_0+1)^{d/2}|K|^{1/2}}\right) e^{-0.5H(X,K))} dK \\
&= \int_{-\infty}^{\infty} \frac{\lambda_0^{d/2}|K|^{(v-d)/2}}{(\lambda_0+1)^{d/2}2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} e^{-0.5H(X,K))} dK \\
&= \int_{-\infty}^{\infty} \frac{\lambda_0^{d/2}|K|^{(v-d)/2}}{(\lambda_0+1)^{d/2}2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} e^{-0.5tr((V^{-1}+XX^T+\mu_0\mu_0^T\lambda_0-\omega\omega^T(\lambda_0+1))K)} dK \\
&= \int_{-\infty}^{\infty} \frac{\lambda_0^{d/2}|K|^{(v-d)/2}}{(\lambda_0+1)^{d/2}2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} e^{-0.5tr(U^{-1}K)} dK \\
&= \frac{\lambda_0^{d/2}}{(\lambda_0+1)^{d/2}2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} \int_{-\infty}^{\infty} |K|^{(v-d)/2} e^{-0.5tr(U^{-1}K)} dK \\
&= \frac{\lambda_0^{d/2}}{(\lambda_0+1)^{d/2}2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} \int_{-\infty}^{\infty} |K|^{((v+1)-d-1)/2} e^{-0.5tr(U^{-1}K)} dK \\
&= \frac{\lambda_0^{d/2}}{(\lambda_0+1)^{d/2}2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} \int_{-\infty}^{\infty} |K|^{(v'-d-1)/2} e^{-0.5tr(U^{-1}K)} dK
\end{aligned}
$$

with

$$U^{-1} = V^{-1} + XX^T + \mu_0\mu_0^T\lambda_0 - \omega\omega^T(\lambda_0 + 1)$$
$$v' = v + 1$$

We can now solve the above integral by recognising it as the integral of a Wishart distribution. The definition of the Wishart distribution is given again below for convenience.

$$P(K) = \frac{|K|^{(v-d-1)/2}}{2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)} e^{-tr(V^{-1}K)/2}$$

$$\begin{aligned}
P(X) =& \frac{\lambda_0^{d/2}}{(\lambda_0+1)^{d/2}2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} \int_{-\infty}^{\infty} |K|^{(v'-d-1)/2} e^{-tr(U^{-1}K)/2} dK \\
=& \frac{\lambda_0^{d/2}}{(\lambda_0+1)^{d/2}2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} \left(2^{v'd/2}|U|^{v'/2}\Gamma_d(\frac{v'}{2})\right) \\
=& \frac{2^{v'd/2}\lambda_0^{d/2}\Gamma_d(\frac{v'}{2})}{(\lambda_0+1)^{d/2}2^{vd/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} \left(|U|^{v'/2}\right) \\
=& \frac{2^{(v+1)d/2}\lambda_0^{d/2}\Gamma_d(\frac{(v+1)}{2})}{2^{vd/2}(\lambda_0+1)^{d/2}|V|^{v/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} \left(|(V^{-1} + XX^T + \mu_0\mu_0^T\lambda_0 - \omega\omega^T(\lambda_0+1))^{-1}|^{(v+1)/2}\right) \\
=& \frac{2^{d/2}\lambda_0^{d/2}\Gamma_d(\frac{(v+1)}{2})}{(\lambda_0+1)^{d/2}\Gamma_d\left(\frac{v}{2}\right)(2\pi)^{d/2}} |V|^{-v/2} \left(|V^{-1} + XX^T + \mu_0\mu_0^T\lambda_0 - \omega\omega^T(\lambda_0+1)|\right)^{-(v+1)/2} \\
=& \frac{\lambda_0^{d/2}\Gamma_d(\frac{(v+1)}{2})}{(\lambda_0+1)^{d/2}\Gamma_d\left(\frac{v}{2}\right)\pi^{d/2}} |V|^{-v/2} \left(|V^{-1} + XX^T + \mu_0\mu_0^T\lambda_0 - \omega\omega^T(\lambda_0+1)|\right)^{-(v+1)/2}
\end{aligned}$$

But, remember:

$$\omega = \frac{X + \lambda_0 \mu_0}{1 + \lambda_0}$$

so

$$\begin{aligned}
P(X) =& \frac{1}{Z} |V|^{-v/2} \left( |V^{-1} + XX^T + \mu_0 \mu_0^T \lambda_0 - \omega \omega^T (\lambda_0 + 1)| \right)^{-(v+1)/2} \\
=& \frac{1}{Z} |V|^{-v/2} \left( |V^{-1} + XX^T + \mu_0 \mu_0^T \lambda_0 - \frac{1}{(1 + \lambda_0)^2} (X + \lambda_0 \mu_0)(X + \lambda_0 \mu_0)^T (\lambda_0 + 1)| \right)^{-(v+1)/2} \\
=& \frac{1}{Z} |V|^{-v/2} \left( |V^{-1} + XX^T + \mu_0 \mu_0^T \lambda_0 - \frac{1}{1 + \lambda_0} (X + \lambda_0 \mu_0)(X + \lambda_0 \mu_0)^T| \right)^{-(v+1)/2} \\
=& \frac{1}{Z} |V|^{-v/2} \left( |V^{-1} + XX^T + \mu_0 \mu_0^T \lambda_0 - \frac{1}{1 + \lambda_0} (XX^T + 2\lambda_0 \mu_0 X^T + \lambda_0^2 \mu_0 \mu_0^T)| \right)^{-(v+1)/2} \\
=& \frac{1}{Z} |V|^{-v/2} \left( |V^{-1} + (\frac{\lambda_0}{1 + \lambda_0}) XX^T + \mu_0 \mu_0^T \lambda_0 - \frac{1}{1 + \lambda_0} (2\lambda_0 \mu_0 X^T + \lambda_0^2 \mu_0 \mu_0^T)| \right)^{-(v+1)/2} \\
=& \frac{1}{Z} |V|^{-v/2} \left( |V^{-1} + \left( \frac{\lambda_0}{1 + \lambda_0} \right) XX^T + \left( \frac{\lambda_0^2 + \lambda_0}{1 + \lambda_0} \right) \mu_0 \mu_0^T - \left( \frac{2\lambda_0}{1 + \lambda_0} \right) \mu_0 X^T - \left( \frac{\lambda_0^2}{1 + \lambda_0} \right) \mu_0 \mu_0^T)| \right)^{-(v+1)/2} \\
=& \frac{1}{Z} |V|^{-v/2} \left( |V^{-1} + \left( \frac{\lambda_0}{1 + \lambda_0} \right) XX^T + \left( \frac{\lambda_0}{1 + \lambda_0} \right) \mu_0 \mu_0^T - \left( \frac{2\lambda_0}{1 + \lambda_0} \right) \mu_0 X^T| \right)^{-(v+1)/2} \\
=& \frac{1}{Z} |V|^{-v/2} \left( |V^{-1} + \left( \frac{\lambda_0}{1 + \lambda_0} \right) \left( XX^T + \mu_0 \mu_0^T - 2\mu_0 X^T \right) | \right)^{-(v+1)/2} \\
=& \frac{1}{Z} |V|^{-v/2} \left( |V^{-1} + \left( \frac{\lambda_0}{1 + \lambda_0} \right) (X - \mu_0)(X - \mu_0)^T| \right)^{-(v+1)/2}
\end{aligned}$$

with

$$\frac{1}{Z} = \frac{\lambda_0^{d/2} \Gamma_d(\frac{(v+1)}{2})}{(\lambda_0 + 1)^{d/2} \Gamma_d \left( \frac{v}{2} \right) \pi^{d/2}}$$

Now, using the matrix determinant lemma:

$$|A + uv^T| = (1 + v^t A^{-1} u)|A|$$

we can rewrite the above as follows

$$
\begin{aligned}
P(X) =& \frac{1}{Z}|V|^{-v/2}\left(|V^{-1}+\left(\frac{\lambda_0}{1+\lambda_0}\right)(X-\mu_0)(X-\mu_0)^T|\right)^{-(v+1)/2}\\
=& \frac{1}{Z}|V|^{-v/2}\left(\left(1+\left(\frac{\lambda_0}{1+\lambda_0}\right)(X-\mu_0)^T V(X-\mu_0)\right)|V^{-1}|\right)^{-(v+1)/2}\\
=& \frac{1}{Z}|V|^{-v/2}|V|^{(v+1)/2}\left(1+(X-\mu_0)^T\left(\frac{\lambda_0 V}{1+\lambda_0}\right)(X-\mu_0)\right)^{-(v+1)/2}\\
=& \frac{1}{Z}|V|^{1/2}\left(1+(X-\mu_0)^T\left(\frac{\lambda_0 V}{1+\lambda_0}\right)(X-\mu_0)\right)^{-(v+1)/2}\\
=& \frac{\lambda_0^{d/2}\Gamma_d(\frac{(v+1)}{2})}{(\lambda_0+1)^{d/2}\Gamma_d\left(\frac{v}{2}\right)\pi^{d/2}}|V|^{1/2}\left(1+(X-\mu_0)^T\left(\frac{\lambda_0 V}{1+\lambda_0}\right)(X-\mu_0)\right)^{-(v+1)/2}\\
=& \frac{\Gamma_d(\frac{(v+1)}{2})}{\Gamma_d\left(\frac{v}{2}\right)\pi^{d/2}}\left|\left(\frac{\lambda_0}{1+\lambda_0}\right)V\right|^{1/2}\left(1+(X-\mu_0)^T\left(\frac{\lambda_0 V}{1+\lambda_0}\right)(X-\mu_0)\right)^{-(v+1)/2}\\
=& \frac{\Gamma_d(\frac{((v-d+1)+d)}{2})\left|\frac{(v-d+1)\lambda_0 V}{1+\lambda_0}\right|^{1/2}}{\Gamma_d\left(\frac{v}{2}\right)(v-d+1)^{d/2}\pi^{d/2}}\left(1+\frac{1}{(v-d+1)}(X-\mu_0)^T\left(\frac{(v-d+1)\lambda_0 V}{1+\lambda_0}\right)(X-\mu_0)\right)^{-(v-d+1)+d)/2}
\end{aligned}
$$

We can know recognise the above expression as a multivariate t- distribution and, by looking at the definition of the distribution (see below), we can determine the parameters of the above multivariate t-distribution in terms of the parameters of our mean and precision priors.

$$
t_{v_X}(X;\mu_X,\Sigma_X=K_X^{-1})=\frac{\Gamma(\frac{v_X+d}{2})|K_X|^{1/2}}{\Gamma(\frac{v_X}{2})v_X^{d/2}\pi^{d/2}}\left(1+\frac{1}{v_X}(X-\mu_X)^T K_X(X-\mu_X)\right)^{-(v_X+d)/2}
$$

The parameters are therefore as follows:

$$
\begin{aligned}
v_X =& v-d+1\\
K_X =& \left(\frac{(v-d+1)\lambda_0}{1+\lambda_0}\right)V
\end{aligned}
$$

But the distribution is not properly normalised: $\Gamma_d(\frac{v}{2})$ in the denominator should be $\Gamma_d(\frac{v-d+1}{2})$ Calculation error?