The Langevin equation for a N-dimensional particle with position $\vec{x} = (x_i), \quad i \in \{1 \dots N\}$ in a harmonic potential centered in $\vec{0}$ is

$$\gamma \frac{\mathrm{d}x_i}{\mathrm{d}t} = -\sum_j \lambda_{ij} x_j + \sqrt{2kT\gamma}\xi_i(t) \tag{1}$$

with $\langle \xi_i(t)\xi_j(t')\rangle = \delta_{ij}\delta(t-t')$. The diffusion coefficient is $D = kT/\gamma$. The force on the particle is represented by the stiffness matrix λ_{ij}. It can be shown that the corresponding Fokker-Planck equation is

$$\gamma \partial_t P = \left(-\sum_{i,j} \frac{\partial}{\partial x_i} \lambda_{ij} x_j + kT \frac{\partial^2}{\partial x_i^2} \right) P \tag{2}$$

The stationnary state solution of this is

$$P_{eq}(\vec{x}) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{C}|}} \exp\left\{ -\frac{1}{2kT} \vec{x}^T \boldsymbol{C}^{-1} \vec{x} \right\} \tag{3}$$

with $\boldsymbol{C} = \lambda^{-1}$. In other words, the potential the particle evolves in is $V(x) = \frac{1}{2}\vec{x}^T \boldsymbol{C}^{-1}\vec{x} = \frac{1}{2}\sum_{i,j} \lambda_{ij} x_i x_j$. For our case, we can set $kT = 1$ for the rest.

Following [1], we can explicitely write the solution of 3. The result is an Ornstein-Uhlenbeck process defined in the following way

$$\begin{aligned} P(\vec{x}) =& \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{C}|}} \exp\left\{ -\frac{1}{2} \vec{x}^T \boldsymbol{C}^{-1} \vec{x} \right\}, \\ P(\vec{x}_2|\vec{x}_1, \Delta t) =& \frac{1}{\sqrt{(2\pi)^N (1 - e^{-2\Delta t}) |\boldsymbol{\Sigma}^{-1}|}} \exp\left\{ -\frac{1}{2} (\vec{x}_2 - \vec{\mu_1})^T \boldsymbol{\Sigma}^{-1} (\vec{x}_2 - \vec{\mu_1}) \right\}. \end{aligned} \tag{4}$$

where

$$\mu_1 = \boldsymbol{\Lambda}\vec{x}_1, \qquad \boldsymbol{\Sigma} = \boldsymbol{C} - \boldsymbol{\Lambda}\boldsymbol{C}\boldsymbol{\Lambda}, \qquad \boldsymbol{\Lambda} = e^{-\gamma \boldsymbol{C}^{-1}\Delta t}.$$

Thus, the average and covariance of variable $\vec{x}_1$ are time dependent through matrix $\boldsymbol{\Lambda}$, which itself depends on the equilibrium properties of the process through $\boldsymbol{C}$, and through the dynamical parameter γ. It is important to notice that $\boldsymbol{\Lambda}$, $\boldsymbol{C}$ and $\boldsymbol{\Sigma}$ all commute and are symetric.

Note: On dimensions. $\boldsymbol{C} \sim [T^2]$, $\boldsymbol{\Lambda} \sim [T^{-2}]$, $\gamma \sim [T^{-1}]$ and $kT \sim [L^2 T^{-2}]$.

What is the probability of observing two configurations $\vec{x}_1$ and $\vec{x}_2$ of the system knowing that they are separated by time Δt? Using the identity $\boldsymbol{C}^{-1} = \boldsymbol{\Sigma}^{-1}(1 - \boldsymbol{\Lambda}^2)$, we find

$$\begin{aligned} \log P(\vec{x}_2, \vec{x}_1, \Delta t) &\propto -\frac{1}{2} \left[\vec{x}_2^T \boldsymbol{\Sigma}^{-1} \vec{x}_2 - 2\vec{x}_2^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Lambda} \vec{x}_1 + 2\vec{x}_1^T \boldsymbol{\Lambda} \boldsymbol{\Sigma}^{-1} \boldsymbol{\Lambda} \vec{x}_1 + \vec{x}_1^T \boldsymbol{C}^{-1} \vec{x}_1 \right] \\ &\propto -\frac{1}{2} \left[\vec{x}_2^T \boldsymbol{\Sigma}^{-1} \vec{x}_2 + \vec{x}_1^T \boldsymbol{\Sigma}^{-1} \vec{x}_1 - 2\vec{x}_1^T (\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}) \vec{x}_2 \right], \end{aligned} \tag{5}$$

which is a gaussian distribution with a block correlation matrix, having $\boldsymbol{\Sigma}^{-1}$ on the diagonal blocks and $-\boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1}$ on the off-diagonal part. Inverting this matrix – like a 2×2 matrix, since everything commutes – yields the following expression for the covariance of configurations $\vec{x}_1$ and $\vec{x}_2$ separated by Δt:

$$\langle \vec{x}_1 \vec{x}_2^T \rangle = \boldsymbol{\Lambda}\boldsymbol{\Sigma}^{-1} \cdot \boldsymbol{\Sigma}^2 (1 - \boldsymbol{\Lambda}^2)^{-1} = \boldsymbol{\Lambda}\boldsymbol{C}. \tag{6}$$

As an exercise, let us compute probability of the smallest non trivial tree, *ie* root $\vec{x}_0$ with children $\vec{x}_1$ and $\vec{x}_2$ and branch length Δt. The probability of observing given configurations on this topology is

$$\begin{aligned} P(\vec{x}_0, \vec{x}_1, \vec{x}_2; \Delta t) &= P(\vec{x}_1|\vec{x}_0) P(\vec{x}_2|\vec{x}_0) P(\vec{x}_0) \\ &\propto \exp -\frac{1}{2} \left\{ \vec{x}_1 \boldsymbol{\Sigma}^{-1} \vec{x}_1 + \vec{x}_2 \boldsymbol{\Sigma}^{-1} \vec{x}_2 - 2(\vec{x}_1 + \vec{x}_2) \boldsymbol{\Sigma}^{-1} \boldsymbol{\Lambda} \vec{x}_0 + \vec{x}_0 \boldsymbol{\Sigma}^{-1} (1 + \boldsymbol{\Lambda}^2) \vec{x}_0 \right\}, \end{aligned} \tag{7}$$

where the identity $\boldsymbol{C}^{-1} = \boldsymbol{\Sigma}^{-1}(1 - \boldsymbol{\Lambda}^2)$ has been used. Integrating this over all values of $\vec{x}_0$ using eq. (8), and remembering that $\boldsymbol{\Sigma}(2\Delta t) = \boldsymbol{C} - \boldsymbol{\Lambda}^2 \boldsymbol{C} \boldsymbol{\Lambda}^2$, we recover equation 5 with $\Delta t \to 2\Delta t$, that is with $\boldsymbol{\Lambda} \to \boldsymbol{\Lambda}^2$.

Note: Gaussian integration

$$\int \exp -\frac{1}{2}\left\{\vec{x}^T A\vec{x} + B^T\vec{x}\right\} \mathrm{d}^n x = \left(\frac{(2\pi)^n}{|A|}\right)^{1/2} \exp\left(\frac{1}{8}B^T A^{-1} B\right) \tag{8}$$

Let us do the same thing for a tree with two levels. Nodes are labelled from 0 to 6, with $\vec{x}_1$ and $\vec{x}_2$ being the children of root $\vec{x}_0$, and so on. The full probability can be written

$$P(\{\vec{x}_i\}_{i=0\ldots6}) = P(\vec{x}_3,\vec{x}_4|\vec{x}_1)P(\vec{x}_5,\vec{x}_6|\vec{x}_2)P(\vec{x}_1|\vec{x}_0)P(\vec{x}_2|\vec{x}_0)P_{eq}(\vec{x}_0) \tag{9}$$

Integrating this over $\vec{x}_0$ gives equation 5 for variables 1 and 2, with $\Delta t \to 2\Delta t$, while the part concering variables 3 to 6 remains untouched. Thus, we have to perform the following integration.

$$\begin{aligned} P(\{\vec{x}_i\}_{i=3\ldots6}) &= \int \mathrm{d}\vec{x}_1\mathrm{d}\vec{x}_2 P(\vec{x}_3,\vec{x}_4|\vec{x}_1)P(\vec{x}_5,\vec{x}_6|\vec{x}_2)\cdot P(\vec{x}_2,\vec{x}_1,2\Delta t) \\ &\propto \int \mathrm{d}\vec{x}_1\mathrm{d}\vec{x}_2 \exp -\frac{1}{2}\left\{(\vec{x}_{3/4}-\boldsymbol{\Lambda}\vec{x}_1)\boldsymbol{C}^{-1}(1-\boldsymbol{\Lambda}^2)^{-1}(\vec{x}_{3/4}-\boldsymbol{\Lambda}\vec{x}_1) + (\vec{x}_{5/6}-\boldsymbol{\Lambda}\vec{x}_2)\boldsymbol{C}^{-1}(1-\boldsymbol{\Lambda}^2)^{-1}(\vec{x}_{5/6}-\boldsymbol{\Lambda}\vec{x}_2)\right. \\ &\qquad \left. +\vec{x}_{1/2}\boldsymbol{C}^{-1}(1-\boldsymbol{\Lambda}^4)^{-1}\vec{x}_{1/2} - 2\vec{x}_1\boldsymbol{C}^{-1}\boldsymbol{\Lambda}^2(1-\boldsymbol{\Lambda}^4)^{-1}\vec{x}_2\right\} \end{aligned} \tag{10}$$

where the notation $\vec{x}_{1/2}$ means that the corresponding term must be repeated for variables 1 and 2. For example, $\vec{x}_{1/2}\boldsymbol{U}\vec{x}_{1/2} = \vec{x}_1\boldsymbol{U}\vec{x}_1 + \vec{x}_2\boldsymbol{U}\vec{x}_2$. Here, we used the fact that when $\Delta t \to 2\Delta t$, $\boldsymbol{\Lambda} \to \boldsymbol{\Lambda}^2$. Since we have to integrate over 1 and 2, let us count linear and quadratic terms in those parameters to apply eq (8). For instance, $\vec{x}_1$:

- Quadratic: $2\boldsymbol{C}^{-1}\boldsymbol{\Lambda}^2(1-\boldsymbol{\Lambda}^2)^{-1} + \boldsymbol{C}^{-1}(1-\boldsymbol{\Lambda}^4)^{-1}$
- Linear: $-2\boldsymbol{C}^{-1}\boldsymbol{\Lambda}(1-\boldsymbol{\Lambda}^2)(\vec{x}_3+\vec{x}_4) - 2\boldsymbol{C}^{-1}\boldsymbol{\Lambda}^2(1-\boldsymbol{\Lambda}^4)^{-1}\vec{x}_2$

From this point, it is quite clear that the result will be gaussian. The output of the integration will give quadratic and cross terms in $\vec{x}_{3/4/5/6}$, with a correlation matrix depending on $\boldsymbol{C}$ and $\boldsymbol{\Lambda}$. The exact expression has to be computed, though.

Let us consider a binary tree with $K+1$ levels, labelled $k \in \{0\ldots K\}$. Nodes at level k are written $X^{\{k\}} = (X^i)_{i=1\ldots2^k}$. In the following, we consider two levels of the tree, k and $k+1$, and introduce the following notation: nodes in level k are written $X^{\{k\}} = (X_0^i)_{i=1\ldots2^k}$ and nodes in $k+1$ $X^{\{k+1\}} = \big((X_1^i, X_2^i)\big)_{i=1\ldots2^k}$, where (X_1^i, X_2^i) are children of node X_0^i.
What we want to know is the probability of a configuration of nodes at level $k+1$, independently of the parent configurations. We assume that this quantity is known for level k, and we try to propagate it down the tree. Furthermore, we assume that $P(X^{\{k\}}) = P(X_0^1\ldots X_0^{2^k})$ is a gaussian. In this case, the probability of observing a given configuration in level $k+1$ is

$$P(X^{\{k+1\}}) = \int \left(\prod_{i=1}^{2^k}\mathrm{d}X_0^i\right) P(X_0^1\ldots X_0^{2^k}) \prod_{i=1}^{2^k} P(X_1^i, X_2^i|X_0^i). \tag{11}$$

Since $P(X_1^i, X_2^i|X_0^i)$ is a gaussian in all three variables, and $P(X_0^1\ldots X_0^{2^k})$ is gaussian by assumption, the resulting distribution for nodes at level $k+1$ will also be a gaussian. And since the root of the tree is by assumption a gaussian variable with correlation matrix $\boldsymbol{C}$, the leaves of the tree are correlated gaussian variable by recursion. Quite clearly, the correlation matrix of the leaves is a combination of $\boldsymbol{C}^{-1}$ and of $\boldsymbol{\Lambda}$. Given the expression found for the 2-levels tree and the 3-levels tree (not finished, but it points in this direction), it is likely that this combination involves a product of $\boldsymbol{C}^{-1}$ with fraction of polynomials of $\boldsymbol{\Lambda}$. Branch length and number of children do not change this, but only the nature of $\boldsymbol{\Lambda}$.
Thus, what would be analytically useful is to propagate the correlation matrix through the tree by finding some recursive equation. In other words, given correlation matrix Ξ_k for level k, find Ξ_{k+1} as a function of Ξ_k and $\boldsymbol{\Lambda}$ by performing the integration in eq (11). Solving this numerically would then give us the likelihood of the data given parameters $\boldsymbol{C}$ in the form of a gaussian function. This could then be inverted by Bayes theorem to obtain an estimator of $\boldsymbol{C}$.

References

[1] Rajesh Singh, Dipanjan Ghosh, and R. Adhikari. Fast bayesian inference of the multivariate ornstein-uhlenbeck process. *arxiv:1706.04961*, 2017.