Supplementary Methods

Modeling expression ranks for noise-tolerant differential expression analysis of scRNA-seq data

Krishan Gupta [1], Manan Lalit [2], Aditya Biswas [3], Chad D. Sanada [4], Cassandra Greene [4], Kyle Hukari [4], Ujjwal Maulik [5], Sanghamitra Bandyopadhyay [6], Naveen Ramalingam [4], Gaurav Ahuja [7], Abhik Ghosh [8*], Debarka Sengupta [1,7,9,10*]

[1] Department of Computer Science and Engineering, Indraprastha Institute of Information Technology, Delhi 110020, India, [2] Max Planck Institute of Molecular Cell Biology and Genetics, Dresden 01307, Germany, [3] Microsoft India Pvt. Ltd., Hyderabad, Telangana 500032, India, [4] Fluidigm Corporation, 2 Tower Place, Suite 2000, South San Francisco, CA 94080, USA [5] Department of Computer Science, Jadavpur University, Kolkata, West Bengal 700032, India, [6] Machine Intelligence Unit, Indian Statistical Institute, Kolkata 700108, India, [7] Department of Computational Biology, Indraprastha Institute of Information Technology, Delhi 110020, India, [8] Interdisciplinary Statistical Research Unit, Indian Statistical Institute, Kolkata 700108, India [9] Centre for Artificial Intelligence, Indraprastha Institute of Information Technology, Delhi 110020, India, [10] Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia.

[*]To whom correspondence should be addressed. Tel: +91 11 2690 7446; Email: debarka@iiitd.ac.in
Correspondence may also be addressed to Abhik Ghosh. Tel: +91 33 2575 2307; Email: abhik.ghosh@isical.ac.in

Derivation of the Fisher Information Matrix

Note that, for our DGBD model with likelihood function given by Eq (2) of the main paper, we have

$$\begin{aligned}
\frac{\partial^2 \text{logL}}{\partial a^2} &= \left(\sum_{r=1}^{r=N} y_r\right) \frac{\partial^2 \log(A)}{\partial a^2} \\
\frac{\partial^2 \text{logL}}{\partial b^2} &= \left(\sum_{r=1}^{r=N} y_r\right) \frac{\partial^2 \log(A)}{\partial b^2} \\
\frac{\partial^2 \text{logL}}{\partial a \partial b} &= \left(\sum_{r=1}^{r=N} y_r\right) \frac{\partial^2 \log(A)}{\partial a \partial b}
\end{aligned} \tag{1}$$

So in order to evaluate the above mentioned double derivatives, the first order derivative $\frac{\partial \text{log A}}{\partial a}$ and $\frac{\partial \text{log A}}{\partial b}$ are determined as follows:

$$\begin{aligned}
\text{log A} &= -\log\left(\sum_{r=1}^{r=N} \frac{(N+1-r)^b}{r^a}\right) \\
\frac{\partial \text{log A}}{\partial a} &= \frac{1}{\left(\sum_{r=1}^{r=N} \frac{(N+1-r)^b}{r^a}\right)} \times \sum_{r=1}^{r=N} \frac{(N+1-r)^b \log r}{r^a} \\
\frac{\partial \text{log A}}{\partial b} &= \frac{-1}{\left(\sum_{r=1}^{r=N} \frac{(N+1-r)^b}{r^a}\right)} \times \sum_{r=1}^{r=N} \frac{(N+1-r)^b \log(N+1-r)}{r^a}
\end{aligned} \tag{2}$$

Re-writing Equation 2 in a more succinct form in the Equation 3 below, we get

$$\frac{\partial \text{log A}}{\partial a} = \frac{u_{1,0}}{u_{0,0}} \quad \text{and} \quad \frac{\partial \text{log A}}{\partial b} = -\frac{u_{0,1}}{u_{0,0}} \tag{3}$$

where $u_{i,j}$s are as defined in the main paper. Evaluating the partial derivatives of $u_{1,0}$, $u_{0,0}$ and $u_{0,1}$ with respect to a and b, in the Equation 4:

$$\begin{aligned}
\frac{\partial u_{1,0}}{\partial a} &= -\sum_{r=1}^{r=N} \frac{(N+1-r)^b (\log r)^2}{r^a} \\
\frac{\partial u_{1,0}}{\partial b} &= \sum_{r=1}^{r=N} \frac{(N+1-r)^b [\log r] [\log (N+1-r)]}{r^a} \\
\frac{\partial u_{0,0}}{\partial a} &= -\sum_{r=1}^{r=N} \frac{(N+1-r)^b \log r}{r^a} \\
\frac{\partial u_{0,0}}{\partial b} &= \sum_{r=1}^{r=N} \frac{(N+1-r)^b \log (N+1-r)}{r^a} \\
\frac{\partial u_{0,1}}{\partial a} &= \sum_{r=1}^{r=N} \frac{(N+1-r)^b [\log r] [\log (N+1-r)]}{r^a} \\
\frac{\partial u_{0,1}}{\partial b} &= -\sum_{r=1}^{r=N} \frac{(N+1-r)^b [\log (N+1-r)]^2}{r^a}
\end{aligned} \tag{4}$$

In a compact form, these can be written more generally, for any $i, j = 0, 1$, as

$$\frac{\partial u_{i,j}}{\partial a} = -u_{i+1,j}, \qquad \frac{\partial u_{i,j}}{\partial b} = u_{i,j+1}. \tag{5}$$

Substituting the above expressions in the formula for Fisher information matrix in Eq (3) of the main paper, we get its simplified form for computation within our ROSeq.