# Smoothed Biasing Forces Yield Unbiased Free Energies with the Extended-System Adaptive Biasing Force Method

Adrien Lesage,[†] Tony Lelièvre,[‡] Gabriel Stoltz,[‡] and Jérôme Hénin[*,†]

†*Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, CNRS, Paris, France*

‡*CERMICS, École des Ponts ParisTech, Marne-la-Vallée, France*

E-mail: jerome.henin@ibpc.fr

Phone: +33 158415166. Fax: +33 158415179

# Appendix: Derivations of formulas used in the text

## CZAR for a scalar variable

Starting from the joint distribution of Equation (**??**):

$$\begin{aligned}
\tilde{\rho}^k(z,\lambda) &= Z^{-1}\rho(z)\exp\left(-\beta k\frac{(\lambda-z)^2}{2}\right)\exp(\beta A^k(\lambda)) \\
\ln\rho(z) &= \ln(Z) + \ln\tilde{\rho}^k(z,\lambda) + \beta k\frac{(\lambda-z)^2}{2} - \beta A^k(\lambda) \\
A(z) &= -\frac{1}{\beta}\ln(Z) - \frac{1}{\beta}\ln\tilde{\rho}^k(z,\lambda) - k\frac{(\lambda-z)^2}{2} + \beta A^k(\lambda)
\end{aligned}$$

Inserting the Bayes formula:

$$\begin{aligned}
A(z) &= -\frac{1}{\beta}\ln(Z) - \frac{1}{\beta}\ln\tilde{\rho}(\lambda|z) - \frac{1}{\beta}\ln\tilde{\rho}(z) - k\frac{(\lambda-z)^2}{2} + \beta A^k(\lambda) \\
A'(z) &= -\frac{1}{\beta}\frac{\mathrm{d}\ln\tilde{\rho}(\lambda|z)}{\mathrm{dz}} - \frac{1}{\beta}\frac{\mathrm{d}\ln\tilde{\rho}(z)}{\mathrm{dz}} + k(\lambda-z) \\
A'(z)\tilde{\rho}(\lambda|z) &= -\frac{1}{\beta}\frac{\mathrm{d}\tilde{\rho}(\lambda|z)}{\mathrm{dz}} - \frac{1}{\beta}\frac{\mathrm{d}\ln\tilde{\rho}(z)}{\mathrm{dz}}\tilde{\rho}(\lambda|z) + k(\lambda-z)\tilde{\rho}(\lambda|z) \\
A'(z)\int\tilde{\rho}(\lambda|z)\mathrm{d}\lambda &= -\frac{1}{\beta}\int\frac{\mathrm{d}\tilde{\rho}(\lambda|z)}{\mathrm{dz}}\mathrm{d}\lambda - \frac{1}{\beta}\frac{\mathrm{d}\ln\tilde{\rho}(z)}{\mathrm{dz}}\int\tilde{\rho}(\lambda|z)\mathrm{d}\lambda + \int k(\lambda-z)\tilde{\rho}(\lambda|z)\mathrm{d}\lambda
\end{aligned}$$

Under appropriate regularity and integrability conditions on the marginal distribution, we swap the integral over $\lambda$ and differentiation with respect to $z$, which leads to $\int\frac{\mathrm{d}\tilde{\rho}(\lambda|z)}{\mathrm{dz}}\mathrm{d}\lambda = 0$. Noting that $\int\tilde{\rho}(\lambda|z)\mathrm{d}\lambda = 1$, we obtain the CZAR expression:

$$A'(z) = -\frac{1}{\beta}\frac{\mathrm{d}\ln\tilde{\rho}(z)}{\mathrm{dz}} + k(\langle\lambda\rangle_z - z).$$

## CZAR for a 2D vector variable

We start from the joint marginal distribution obtained as the $\xi$-conditioned integral of the partition function, as in Equation (**??**):

$$\tilde{\rho}^k(z_1,\lambda_1,z_2,\lambda_2) = Z^{-1}\tilde{\rho}(z_1,z_2)\exp\left(-\beta k_1\frac{(\lambda_1-z_1)^2}{2}\right)\exp\left(-\beta k_2\frac{(\lambda_2-z_2)^2}{2}\right)\exp(\beta A^k(\lambda_1,\lambda_2))$$

$$A(z_1,z_2) = -\frac{1}{\beta}\ln Z - \frac{1}{\beta}\ln\tilde{\rho}^k(z_1,\lambda_1,z_2,\lambda_2) - k_1\frac{(z_1-\lambda_1)^2}{2} - k_2\frac{(z_2-\lambda_2)^2}{2} + A^k(\lambda_1,\lambda_2)$$

Inserting the Bayes formula as above:

$$\begin{aligned} A(z_1,z_2) &= -\frac{1}{\beta}\ln Z - \frac{1}{\beta}\ln\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2) - \frac{1}{\beta}\ln\tilde{\rho}(z_1,z_2) - k_1\frac{(z_1-\lambda_1)^2}{2} \\ &\quad - k_2\frac{(z_2-\lambda_2)^2}{2} + A^k(\lambda_1,\lambda_2) \\ \frac{\partial A(z_1,z_2)}{\partial z_1} &= -\frac{1}{\beta}\frac{\partial\ln\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2)}{\partial z_1} - \frac{1}{\beta}\frac{\partial\ln\tilde{\rho}(z_1,z_2)}{\partial z_1} - k_1(z_1-\lambda_1) \end{aligned}$$

$$\begin{aligned} \frac{\partial A(z_1,z_2)}{\partial z_1}\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2) &= -\frac{1}{\beta}\frac{\partial\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2)}{\partial z_1} - \frac{1}{\beta}\frac{\partial\ln\tilde{\rho}(z_1,z_2)}{\partial z_1}\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2) \\ &\quad - k_1(z_1-\lambda_1)\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2) \\ \frac{\partial A(z_1,z_2)}{\partial z_1}\int\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2)\mathrm{d}\lambda_1\mathrm{d}\lambda_2 &= -\frac{1}{\beta}\int\frac{\partial\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2)}{\partial z_1}\mathrm{d}\lambda_1\mathrm{d}\lambda_2 \\ &\quad - \frac{1}{\beta}\frac{\partial\ln\tilde{\rho}(z_1,z_2)}{\partial z_1}\int\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2)\mathrm{d}\lambda_1\mathrm{d}\lambda_2 \\ &\quad - k_1\int(z_1-\lambda_1)\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2)\mathrm{d}\lambda_1\mathrm{d}\lambda_2 \end{aligned}$$

Under the same regularity assumptions as in the scalar case, the first RHS term is zero. Noting that the integral of the joint conditional distribution $\tilde{\rho}(\lambda_1,\lambda_2|z_1,z_2)$ is unity, we obtain the two-dimensional CZAR expression:

$$\frac{\partial A(z_1,z_2)}{\partial z_1} = -\frac{1}{\beta}\frac{\partial\ln\tilde{\rho}(z_1,z_2)}{\partial z_1} + k_1(\langle\lambda_1\rangle_{z_1,z_2} - z_1).$$

## Distance between a function and its Gaussian convolution

Given a function $f$ and $G_\sigma$ the Gaussian kernel of variance $\sigma^2$, we consider

$$(G_\sigma * f)(x) \equiv \frac{1}{\sqrt{2\pi}\sigma} \int \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) f(y)dy$$

as an approximation of $f(x)$ when $\sigma$ tends towards zero.

The local error of this estimate is:

$$\begin{aligned}
|(G_\sigma * f)(x) - f(x)| &= \left| \frac{1}{\sqrt{2\pi}\sigma} \int \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) f(y)dy - f(x) \right| \\
&= \left| \frac{1}{\sqrt{2\pi}\sigma} \int \exp\left(-\frac{y^2}{2\sigma^2}\right) f(x-y)dy - f(x) \right| \\
&= \left| \frac{1}{\sqrt{2\pi}\sigma} \int \exp\left(-\frac{y^2}{2\sigma^2}\right) (f(x-y) - f(x))dy \right| \\
&= \left| \frac{1}{\sqrt{2\pi}} \int \exp(-z^2/2)(f(x-\sigma z) - f(x))dz \right|,
\end{aligned}$$

where we substitute $z \equiv y/\sigma$ in the last line.

Assuming that the second derivative of $f$ is bounded, we can write the following Taylor expansion:

$$f(x - \sigma z) = f(x) - f'(x)\sigma z + f''(x - \theta_{x,z}\sigma z)\frac{(\sigma z)^2}{2}$$

where $\theta_{x,z} \in [0, 1]$. Hence

$$\begin{aligned}
|(G_\sigma * f)(x) - f(x)| &= \left| 1/\sqrt{2\pi} \int \exp(-z^2/2) \left( -f'(x)\sigma z + f''(x - \theta_{x,z}\sigma z)\frac{(\sigma z)^2}{2} \right) dz \right| \\
&= \left| 1/\sqrt{2\pi} \int \exp(-z^2/2) f''(x - \theta_{x,z}\sigma z)\frac{(\sigma z)^2}{2} dz \right| \\
&\le ||f''||_{L^\infty} \left| 1/\sqrt{2\pi} \int \exp(-z^2/2)\frac{(\sigma z)^2}{2} dz \right|,
\end{aligned}$$

where $||f''||_{L^\infty}$ is the essential supremum of $f''$, i.e. the smallest real number greater than $f''(x)$

for almost all $x$ (up to a set of measure zero). Finally, we obtain:

$$|(G_\sigma * f)(x) - f(x)| \leq ||f''||_{L^\infty} \frac{\sigma^2}{2}.$$