

## GROMACS - Bug #2487

### AWH covering detection delayed in certain cases

04/25/2018 11:48 AM - Viveca Lindahl

<b>Status:</b>	Closed	<b>Difficulty:</b>	uncategorized
<b>Priority:</b>	Normal		
<b>Assignee:</b>	Viveca Lindahl		
<b>Category:</b>	mdrun		
<b>Target version:</b>	2018.2		
<b>Affected version - extra info:</b>			
<b>Affected version:</b>	2018.1		

#### Description

For a Brownian dynamics system with a 2d reaction coordinate, the AWH detection of covering gets delayed because the check frequency is determined to be too high. The check frequency is set to scale with the number of points in the grid but the covering time should not directly depend on the discretization. For Brownian dynamics, the time step can be chosen to be large relative to the diffusion time. This exposed the problem since the number of number of samples needed per covering time is relatively small.

#### Associated revisions

##### Revision 26db2acf - 06/12/2018 12:02 PM - Viveca Lindahl

Changed histogram check interval for multidimensional AWH bias.

The step interval for performing covering checks depended explicitly on the total number of points in the AWH histogram. The covering checks themselves however only depends directly on the extent of each one-dimensional axis. For a multidimensional grid in combination with relatively few samples to cover the sampling interval (e.g. using Brownian dynamics), this could lead to a delayed detection of the covering.

The same step interval was also used for checking for histogram anomalies, which however only generates warnings and is not integral to the AWH method itself.

Now the covering check interval is instead determined from the number of sigmas, the "width" of one sample, required to cover each dimension. Since there is about 1 point per sigma, the dependency of the check interval on the number of points is essentially unchanged in the one-dimensional case. However, this relation is not numerically exact. Therefore, this change also requires updating reference data for a regression test to be updated.

The check interval for histogram anomalies is set to be the same as for the covering, but could in the future be made less frequent.

Added release note.

Fixes #2487

Change-Id: I5af0d48436664d8fdfe8bafa05cde3cdae27e45a

#### History

##### #1 - 04/25/2018 11:56 AM - Gerrit Code Review Bot

Gerrit received a related patchset '2' for Issue [#2487](#).

Uploader: Viveca Lindahl ([vivecalindahl@gmail.com](mailto:vivecalindahl@gmail.com))

Change-Id: gromacs~release-2018~I5af0d48436664d8fdfe8bafa05cde3cdae27e45a

Gerrit URL: <https://gerrit.gromacs.org/7790>

##### #2 - 04/26/2018 10:47 AM - Berk Hess

The covering frequency does not depend on discretization. The discretization is set by the force constant. So the covering frequency depends on force constant, but that makes sense, I think, because this changes the requested resolution.

**#3 - 04/26/2018 11:39 AM - Berk Hess**

Looking at the threshold value for covering, I think the least amount of samples you need to cover  $\#points\_in\_dim/\sqrt{2\pi}$ . Or did I miss something? So the fastest you can cover is in  $0.4*\#points\_in\_dim$ . So we check somewhat infrequent for the fastest possible covering case. But it's rather unlikely that diffusion covers the range homogeneously. The time should go up with  $\sqrt{\#points}$ .

**#4 - 04/26/2018 11:55 AM - Viveca Lindahl**

Berk Hess wrote:

The covering frequency does not depend on discretization. The discretization is set by the force constant. So the covering frequency depends on force constant, but that makes sense, I think, because this changes the requested resolution.

There's a negation missing somewhere? The covering frequency should not depend on either force constant or discretization (the latter two are functions of each other).

Berk Hess wrote:

Looking at the threshold value for covering, I think the least amount of samples you need to cover  $\#points\_in\_dim/\sqrt{2\pi}$ . Or did I miss something? So the fastest you can cover is in  $0.4*\#points\_in\_dim$ . So we check somewhat infrequent for the fastest possible covering case. But it's rather unlikely that diffusion covers the range homogeneously. The time should go up with  $\sqrt{\#points}$ .

The covering criterion should AFAIR not depend on the # of points. The weight required for each point will decrease with the number of points. Correct me if I'm wrong...

I did realize however that the problem is there are two quite different checks being determined by the `isCheckStep` function. I'm adding another change that splits this into two: one for covering and one for histogram anomalies. After looking closer the latter actually should depend on the number of points, since there the realized visits histogram, which is a simple binning histogram, is being compared to the weight histogram (to see e.g. if the force constant may be too low).

**#5 - 04/26/2018 12:16 PM - Berk Hess**

I don't see how the threshold depends on the number of points. It looks like the prefactor in the Gaussian to me.

**#6 - 04/26/2018 12:21 PM - Viveca Lindahl**

Berk Hess wrote:

I don't see how the threshold depends on the number of points. It looks like the prefactor in the Gaussian to me.

exactly, it should not depend on the number of points in the end

```
weightThreshold *= grid.axis(d).spacing()*std::sqrt(dimParams[d].betak*0.5*M_1_PI);
```

the spacing will decrease with increasing number of points.

**#7 - 04/26/2018 01:29 PM - Berk Hess**

`spacing=1/sqrt(betak)`  
so the threshold is fixed at 0.4 and does not change with the number of points.

**#8 - 04/26/2018 01:56 PM - Viveca Lindahl**

Berk Hess wrote:

`spacing=1/sqrt(betak)`  
so the threshold is fixed at 0.4 and does not change with the number of points.

You are right. The confusion on my part comes from that the transition of the coordinate should not depend on discretization, but because we don't sample every step and the covering criterion checks for covering each point rather than transitions, there's a built in dependency on the number of points.

**#9 - 06/12/2018 03:15 PM - Anonymous**

- Status changed from New to Resolved

Applied in changeset [26db2acf4a3a0e5dbe3be986151c1e6bdd7ce1a1](#).

**#10 - 06/12/2018 03:16 PM - Mark Abraham**

Is this issue resolved by the recently submitted changes?

**#11 - 06/14/2018 10:50 AM - Mark Abraham**

*- Status changed from Resolved to Closed*

**Files**

---

bd-2d.tgz	2.91 KB	04/25/2018	Viveca Lindahl
-----------	---------	------------	----------------